```
                      GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         December 29, 2004, 13:09:29 ; Search time 24 Seconds
                                           (without alignments)
                                           19.343 Million cell updates/sec

Title:          US-09-606-910E-3
Perfect score:  41
Sequence:       1 PDTRPAP 7

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters:      478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                2:  /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                3:  /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                4:  /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                5:  /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                6:  /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES
                   %
Result           Query
  No.   Score    Match Length DB  ID                      Description
 -----------------------------------------------------------------------
    1      41    100.0      8  4  US-09-641-833-2          Sequence 2, Appli
    2      41    100.0      9  2  US-08-288-059-13         Sequence 13, Appl
    3      41    100.0      9  2  US-08-288-059-27         Sequence 27, Appl
    4      41    100.0      9  2  US-08-666-473-110        Sequence 110, App
    5      41    100.0      9  3  US-08-134-198E-42        Sequence 42, Appl
    6      41    100.0      9  3  US-09-043-731-9          Sequence 9, Appli
    7      41    100.0      9  3  US-09-043-731-10         Sequence 10, Appl
    8      41    100.0      9  4  US-09-593-870A-8         Sequence 8, Appli
    9      41    100.0      9  4  US-09-593-870A-29        Sequence 29, Appl
   10      41    100.0      9  4  US-09-593-870A-69        Sequence 69, Appl
   11      41    100.0      9  4  US-09-497-232-2          Sequence 2, Appli
   12      41    100.0      9  4  US-09-497-232-19         Sequence 19, Appl
   13      41    100.0      9  4  US-09-497-232-20         Sequence 20, Appl
   14      41    100.0     10  3  US-09-339-944-9          Sequence 9, Appli
   15      41    100.0     10  4  US-09-497-232-28         Sequence 28, Appl
   16      41    100.0     10  4  US-09-497-232-29         Sequence 29, Appl
   17      41    100.0     10  4  US-09-497-232-30         Sequence 30, Appl
   18      41    100.0     10  4  US-09-497-232-31         Sequence 31, Appl
   19      41    100.0     10  4  US-09-651-265-9          Sequence 9, Appli
   20      41    100.0     13  4  US-09-914-259-77         Sequence 77, Appl
   21      41    100.0     14  2  US-08-463-230A-12        Sequence 12, Appl
   22      41    100.0     15  1  US-08-122-546-18         Sequence 18, Appl
   23      41    100.0     15  2  US-08-764-938-18         Sequence 18, Appl
   24      41    100.0     15  3  US-09-131-052-18         Sequence 18, Appl
   25      41    100.0     15  3  US-09-131-053A-18        Sequence 18, Appl
   26      41    100.0     16  3  US-09-339-944-7          Sequence 7, Appli
   27      41    100.0     16  3  US-08-737-896-1          Sequence 1, Appli
```

```
   28      41    100.0     16  3  US-09-043-731-18         Sequence 18, Appl
   29      41    100.0     16  4  US-09-497-232-15         Sequence 15, Appl
   30      41    100.0     16  4  US-09-651-265-12         Sequence 12, Appl
   31      41    100.0     16  5  PCT-US96-09951-1         Sequence 1, Appli
   32      41    100.0     17  3  US-08-766-521-2          Sequence 2, Appli
   33      41    100.0     17  3  US-08-755-413-2          Sequence 2, Appli
   34      41    100.0     19  1  US-08-099-354-3          Sequence 3, Appli
   35      41    100.0     19  1  US-08-099-354-5          Sequence 5, Appli
   36      41    100.0     19  2  US-08-288-059-9          Sequence 9, Appli
   37      41    100.0     19  2  US-08-288-059-11         Sequence 11, Appl
   38      41    100.0     19  4  US-09-593-870A-22        Sequence 22, Appl
   39      41    100.0     20  1  US-08-328-536-1          Sequence 1, Appli
   40      41    100.0     20  2  US-08-288-059-1          Sequence 1, Appli
   41      41    100.0     20  2  US-08-288-059-12         Sequence 12, Appl
   42      41    100.0     20  2  US-08-288-059-32         Sequence 32, Appl
   43      41    100.0     20  2  US-08-902-516-20         Sequence 20, Appl
   44      41    100.0     20  2  US-08-833-807-1          Sequence 1, Appli
   45      41    100.0     20  2  US-08-833-807-8          Sequence 8, Appli

                                ALIGNMENTS


RESULT 1
US-09-641-833-2
; Sequence 2, Application US/09641833
; Patent No. 6716966
; GENERAL INFORMATION:
;  APPLICANT: Madiyalakan, Ragupathy
;  TITLE OF INVENTION: Therapeutic Binding Agents Against MUC-1 Antigen and
;  TITLE OF INVENTION:  Methods
;  TITLE OF INVENTION:  for Their Use
;  FILE REFERENCE: 107823.127
;  CURRENT APPLICATION NUMBER: US/09/641,833
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 8
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: MUC-1 antigen
US-09-641-833-2

  Query Match            100.0%;  Score 41;  DB 4;  Length 8;
  Best Local Similarity  100.0%;  Pred. No. 3.8e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          2 PDTRPAP 8


RESULT 2
US-08-288-059-13
; Sequence 13, Application US/08288059
; Patent No. 5827666
;  GENERAL INFORMATION:
;    APPLICANT:  FINN, OLIVERA J.
;    APPLICANT:  FONTENOT, J. D.
;    APPLICANT:  MONTELARO, RONALD C.
;    TITLE OF INVENTION:  SYNTHETIC MULTIPLE TANDEM REPEAT MUCIN
;    TITLE OF INVENTION:  AND MUCIN-LIKE PEPTIDES, AND USES THEREOF
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN DARBY & CUSHMAN, L.L.P.
;      STREET:  1100 NEW YORK AVENUE, N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/288,059
;      FILING DATE:  08-AUG-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CHAPIN, MARLANA K.
;      REGISTRATION NUMBER:  35,843
;      REFERENCE/DOCKET NUMBER:  61137/205204
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-861-3711
;      TELEFAX:  202-822-0944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-288-059-13

  Query Match             100.0%;  Score 41;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          1 PDTRPAP 7


RESULT 3
US-08-288-059-27
; Sequence 27, Application US/08288059
; Patent No. 5827666
;  GENERAL INFORMATION:
;    APPLICANT:  FINN, OLIVERA J.
;    APPLICANT:  FONTENOT, J. D.
;    APPLICANT:  MONTELARO, RONALD C.
;    TITLE OF INVENTION:  SYNTHETIC MULTIPLE TANDEM REPEAT MUCIN
;    TITLE OF INVENTION:  AND MUCIN-LIKE PEPTIDES, AND USES THEREOF
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN DARBY & CUSHMAN, L.L.P.
;      STREET:  1100 NEW YORK AVENUE, N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/288,059
;      FILING DATE:  08-AUG-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CHAPIN, MARLANA K.
;      REGISTRATION NUMBER:  35,843
;      REFERENCE/DOCKET NUMBER:  61137/205204
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-861-3711
;      TELEFAX:  202-822-0944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
```

```
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-288-059-27

  Query Match             100.0%;  Score 41;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          2 PDTRPAP 8


RESULT 4
US-08-666-473-110
; Sequence 110, Application US/08666473
; Patent No. 5843713
;  GENERAL INFORMATION:
;    APPLICANT:  YOSHIDA, Aruto
;    APPLICANT:  TAKEUCHI, Makoto
;    TITLE OF INVENTION:  PEPTIDE SEQUENCE THAT FORMS MUCIN SUGAR
;    TITLE OF INVENTION:  CHAIN AND TECHNIQUE FOR MODIFYING PROTEIN TO BE LINKED
;    TITLE OF INVENTION:  WITH MUCIN SUGAR CHAIN
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/666,473
;      FILING DATE:  19-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/JP95/02238
;      FILING DATE:  01-NOV-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-22101
;      FILING DATE:  09-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-269111
;      FILING DATE:  01-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16887/837
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  110:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-666-473-110

  Query Match             100.0%;  Score 41;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          1 PDTRPAP 7
```

```
                        GenCore version 5.1.6
              Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         December 29, 2004, 15:36:58 ; Search time 2046 Seconds
                                         (without alignments)
                                         124.672 Million cell updates/sec

Title:          US-09-606-910E-3
Perfect score:  41
Sequence:       1 PDTRPAP 7

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters:      158194

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09606910/runat_29122004_131140_4695/app_query.fasta_1.199
-DB=EST -QFMT=fastap -SUFFIX=50nt.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=50
-USER=US09606910_@CGN_1_1_3437_@runat_29122004_131140_4695 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      EST:*
                1:  gb_est1:*
                2:  gb_est2:*
                3:  gb_htc:*
                4:  gb_est3:*
                5:  gb_est4:*
                6:  gb_est5:*
                7:  gb_est6:*
                8:  gb_gss1:*
                9:  gb_gss2:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result          Query
  No.   Score   Match Length DB  ID                     Description
----------------------------------------------------------------------------
     1     34    82.9     31  1  AI682900               AI682900 wc69g04.x
c    2     34    82.9     32  6  CA585831               CA585831 LBA00273.
     3     31    75.6     31  9  TA83D08Q               AL462389 T. brucei
c    4     31    75.6     39  8  AZ834606               AZ834606 2M0117G22
     5     30    73.2     22  1  AI357851               AI357851 qv73g11.x
c    6     30    73.2     36  4  BG122441               BG122441 602353382
c    7     30    73.2     50  1  AU105505               AU105505 AU105505
c    8     29    70.7     24  8  AZ936903               AZ936903 2M0193E20
     9     29    70.7     40  9  BX149818               BX149818 Danio rer
c   10     29    70.7     41  8  AZ937325               AZ937325 2M0195A02
    11     29    70.7     46  1  AI284041               AI284041 qt72d04.x
    12     29    70.7     48  9  BX949286               BX949286 Arabidops
c   13     29    70.7     50  1  AU105698               AU105698 AU105698
c   14     29    70.7     50  1  AU105699               AU105699 AU105699
c   15     28    68.3     25  1  AI521520               AI521520 to64b05.x
c   16     28    68.3     27  8  AZ835367               AZ835367 2M0129M03
c   17     28    68.3     32  8  AZ764538               AZ764538 1M0560N21
    18     28    68.3     33  4  BI822372               BI822372 603037932
    19     28    68.3     37  1  AI689454               AI689454 tx94c06.x
    20     28    68.3     40  1  AA934396               AA934396 oo49b02.s
    21     28    68.3     40  8  CC060546               CC060546 EY04640-5
c   22     28    68.3     44  8  AZ498888               AZ498888 1M0336E21
    23     28    68.3     45  1  AV833912               AV833912 AV833912
c   24     28    68.3     46  1  AI098710               AI098710 vg80b09.r
c   25     28    68.3     50  1  AU104946               AU104946 AU104946
    26     28    68.3     50  9  CR138599               CR138599 Forward s
c   27     28    68.3     50  9  CC941568               CC941568 02S2034-0
c   28     28    68.3     50  9  CC941572               CC941572 02S2034-0
    29     27    65.9     23  9  TA134E10Q              AL466333 T. brucei
    30     27    65.9     28  1  AI358659               AI358659 qx60e07.x
c   31     27    65.9     31  4  BG870909               BG870909 602792332
c   32     27    65.9     33  1  AL588927               AL588927 AL588927
    33     27    65.9     33  4  BI684367               BI684367 603308213
    34     27    65.9     34  1  AI224571               AI224571 qw96a07.x
c   35     27    65.9     37  1  AA906610               AA906610 ok79g04.s
    36     27    65.9     38  9  DR32H3S                AL978781 Danio rer
    37     27    65.9     39  9  TA285C10P              AL487756 T. brucei
    38     27    65.9     41  8  BH866295               BH866295 SALK_1011
    39     27    65.9     46  9  BX171949               BX171949 Danio rer
c   40     27    65.9     47  8  AZ331536               AZ331536 1M0059H04
    41     27    65.9     48  2  BF507210               BF507210 3659P-23
c   42     27    65.9     49  2  BF203474               BF203474 601865731
c   43     27    65.9     49  8  BH626797               BH626797 1007065C0
    44     27    65.9     50  1  AA633761               AA633761 ac27c07.s
    45     27    65.9     50  1  AU103440               AU103440 AU103440

                              ALIGNMENTS

RESULT 1
AI682900
LOCUS       AI682900                31 bp    mRNA    linear   EST 26-MAY-1999
DEFINITION  wc69g04.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2323926 3'
            similar to WP:C34D4.11 CE17505 ;, mRNA sequence.
ACCESSION   AI682900
VERSION     AI682900.1  GI:4893082
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Life Technologies catalog #: 11548-013
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1..31
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
```

```
                    /clone="IMAGE:2323926"
                    /tissue_type="adenocarcinoma"
                    /lab_host="DH10B"
                    /clone_lib="NCI_CGAP_Pan1"
                    /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
                    Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                    Average insert size 1.72 kb. Life Technologies catalog #:
                    11548-013"
ORIGIN

Alignment Scores:
Pred. No.:              3.83e+03        Length:         31
Score:                  34.00           Matches:        6
Percent Similarity:     85.71%          Conservative:   0
Best Local Similarity:  85.71%          Mismatches:     1
Query Match:            82.93%          Indels:         0
DB:                     1               Gaps:           0

US-09-606-910E-3 (1-7) x AI682900 (1-31)

Qy          1 ProAspThrArgProAlaPro 7
              |||   |||||||||||||||
Db          1 CCCCCCACGCGCCCCGCCCCC 21

RESULT 2
CA585831/c
LOCUS       CA585831                  32 bp    mRNA    linear   EST 12-JAN-2004
DEFINITION  LBA00273.BOST4 cDNA from mouse aorta Mus musculus cDNA, mRNA
            sequence.
ACCESSION   CA585831
VERSION     CA585831.1  GI:40790953
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 32)
  AUTHORS   Borang,S., Andersson,T., Thelin,A., Odeberg,J. and Lundeberg,J.
  TITLE     Vascular gene expression in atherosclerotic plaque prone regions
            analysed by representational difference analysis
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Andersson Tove
            Department of Biotechnology
            KTH
            Teknikringen 34, plan 6, 100 44 Stockholm, Sweden
            Tel: +46 8 790 71 29
            Fax: +46 8 245452
            Email: tove@biochem.kth.se
            Genes upregulated in plaque prone regions (DP1 and DP2)
            Seq primer: CTA TGA CCA TGA TTA CGC CAA G.
FEATURES             Location/Qualifiers
     source          1..32
                    /organism="Mus musculus"
                    /mol_type="mRNA"
                    /strain="ApoE-/- and LDLR-/- on C57BL/6x 129 background"
                    /db_xref="taxon:10090"
                    /sex="male"
                    /dev_stage="8 weeks old"
                    /clone_lib="cDNA from mouse aorta"
                    /note="Organ: aorta; Site_1: DpnII; Site_2: DpnII; CDNA
                    was prepared from whole aorta divided in atherosclerotic
                    plaque prone regions (aortic arch and abdominal aorta
                    proximal part) and less plaque prone regions (descending
                    thoracic aorta and abdominal aorta distal part). CDNA was
                    fragmented with DpnII, linker ligated and amplified to
                    generated starting material for representational
                    difference analysis (RDA). The two cDNA pools were
                    subjected to iterative RDA subtraction and amplification
                    to enrich for gene fragments differentially expressed at
                    early stages of atherosclerosis."
ORIGIN

Alignment Scores:
```

```
Pred. No.:              3.94e+03        Length:         32
Score:                  34.00           Matches:        5
Percent Similarity:     100.00%         Conservative:   2
Best Local Similarity:  71.43%          Mismatches:     0
Query Match:            82.93%          Indels:         0
DB:                     6               Gaps:           0

US-09-606-910E-3 (1-7) x CA585831 (1-32)

Qy          1 ProAspThrArgProAlaPro 7
              ||||||::::::|||||||||
Db         23 CCTGACTCTAAGCCAGCCCCT 3

RESULT 3
TA83D08Q
LOCUS       TA83D08Q                  31 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 83d08, reverse sequence,
            genomic survey sequence.
ACCESSION   AL462389
VERSION     AL462389.1  GI:11861024
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nhl@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1..31
                     /organism="Trypanosoma brucei"
                     /mol_type="genomic DNA"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="83d08"
ORIGIN

Alignment Scores:
Pred. No.:              1.25e+04        Length:         31
Score:                  31.00           Matches:        5
Percent Similarity:     71.43%          Conservative:   0
Best Local Similarity:  71.43%          Mismatches:     2
Query Match:            75.61%          Indels:         0
DB:                     9               Gaps:           0

US-09-606-910E-3 (1-7) x TA83D08Q (1-31)

Qy          1 ProAspThrArgProAlaPro 7
              |||   |||||||||   |||
Db          1 CCCAGTACTCGACCCACTCCC 21

RESULT 4
AZ834606/c
LOCUS       AZ834606                  39 bp    DNA     linear   GSS 20-FEB-2001
DEFINITION  2M0117G22F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         December 29, 2004, 17:05:04 ; Search time 360 Seconds
                                          (without alignments)
                                          109.827 Million cell updates/sec

Title:          US-09-606-910E-3
Perfect score:  41
Sequence:       1 PDTRPAP 7

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       4176236 seqs, 2824127955 residues

Total number of hits satisfying chosen parameters:      2810194

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09606910/runat_29122004_131142_4797/app_query.fasta_1.199
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=50nt.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=50 -USER=US09606910_@CGN_1_1_480_@runat_29122004_131142_4797 -NCPU=6
-ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA:*
                1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
                10:  /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
                12:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
                13:  /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
                14:  /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
                15:  /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
                16:  /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
                17:  /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
                18:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
                19:  /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
                20:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
                21:  /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
            %
```

```
Result          Query
  No.   Score   Match Length DB  ID                          Description
-------------------------------------------------------------------------------
c   1     32    78.0     25  16  US-10-407-449-6              Sequence 6, Appli
c   2     32    78.0     25  17  US-10-660-897-6              Sequence 6, Appli
    3     32    78.0     30  15  US-10-273-180-19             Sequence 19, Appl
    4     32    78.0     39  9  US-09-894-633A-37            Sequence 37, Appl
    5     32    78.0     47  10  US-09-993-346-74             Sequence 74, Appl
    6     32    78.0     50  10  US-09-993-346-95             Sequence 95, Appl
c   7     31    75.6     25  15  US-10-098-263B-47201         Sequence 47201, A
c   8     31    75.6     25  15  US-10-098-263B-62272         Sequence 62272, A
c   9     31    75.6     25  15  US-10-098-263B-105312        Sequence 105312,
   10     31    75.6     25  15  US-10-098-263B-110205        Sequence 110205,
   11     31    75.6     25  15  US-10-098-263B-110206        Sequence 110206,
   12     31    75.6     41  16  US-10-035-833A-96            Sequence 96, Appl
c  13     31    75.6     41  16  US-10-035-833A-251           Sequence 251, App
c  14     31    75.6     41  16  US-10-035-833A-2337          Sequence 2337, Ap
c  15     31    75.6     41  16  US-10-035-833A-3742          Sequence 3742, Ap
   16     31    75.6     41  16  US-10-035-833A-5390          Sequence 5390, Ap
c  17     31    75.6     41  16  US-10-035-833A-6720          Sequence 6720, Ap
c  18     30    73.2     17  15  US-10-238-700-3511           Sequence 3511, Ap
c  19     30    73.2     20  14  US-10-044-692-225            Sequence 225, App
c  20     30    73.2     20  15  US-10-044-539-225            Sequence 225, App
c  21     30    73.2     20  15  US-10-325-810-458            Sequence 458, App
   22     30    73.2     20  17  US-10-688-706-291            Sequence 291, App
   23     30    73.2     20  17  US-10-688-706-356            Sequence 356, App
   24     30    73.2     20  17  US-10-688-706-456            Sequence 456, App
c  25     30    73.2     20  18  US-10-877-124-458            Sequence 458, App
c  26     30    73.2     20  18  US-10-877-022-458            Sequence 458, App
   27     30    73.2     25  14  US-10-044-692-208            Sequence 208, App
   28     30    73.2     25  15  US-10-044-539-208            Sequence 208, App
   29     30    73.2     25  15  US-10-098-263B-4001          Sequence 4001, Ap
c  30     30    73.2     25  15  US-10-098-263B-5668          Sequence 5668, Ap
c  31     30    73.2     25  15  US-10-098-263B-16311         Sequence 16311, A
c  32     30    73.2     25  15  US-10-098-263B-43395         Sequence 43395, A
   33     30    73.2     25  15  US-10-098-263B-108922        Sequence 108922,
c  34     30    73.2     25  15  US-10-285-976-98             Sequence 98, Appl
   35     30    73.2     25  15  US-10-325-810-441            Sequence 441, App
   36     30    73.2     25  18  US-10-877-124-441            Sequence 441, App
   37     30    73.2     25  18  US-10-877-022-441            Sequence 441, App
c  38     30    73.2     26  16  US-10-465-811-111            Sequence 111, App
   39     30    73.2     26  16  US-10-460-754-9              Sequence 9, Appli
c  40     30    73.2     28  16  US-10-465-811-112            Sequence 112, App
c  41     30    73.2     30  16  US-10-465-811-122            Sequence 122, App
c  42     30    73.2     30  16  US-10-465-811-123            Sequence 123, App
c  43     30    73.2     31  9  US-09-733-042-37             Sequence 37, Appl
   44     30    73.2     40  16  US-10-447-839A-78            Sequence 78, Appl
   45     30    73.2     40  16  US-10-447-839A-93            Sequence 93, Appl

                              ALIGNMENTS

RESULT 1
US-10-407-449-6/c
; Sequence 6, Application US/10407449
; Publication No. US20040005601A1
; GENERAL INFORMATION:
;  APPLICANT: Siddiqui-Jain, Adam
;  APPLICANT:  Hurley, Laurence
;  APPLICANT:  Farrell, Thomas
;  APPLICANT:  Grand, Cory
;  APPLICANT:  Bearss, David
;  TITLE OF INVENTION: METHODS FOR TARGETING QUADRUPLEX DNA
;  FILE REFERENCE: 53223-20004.00
;  CURRENT APPLICATION NUMBER: US/10/407,449
;  CURRENT FILING DATE:  2003-04-04
;  PRIOR APPLICATION NUMBER: US 60/404,966
;  PRIOR FILING DATE: 2002-08-04
;  PRIOR APPLICATION NUMBER: US 60/370,358
;  PRIOR FILING DATE: 2002-04-05
;  PRIOR APPLICATION NUMBER: Unknown
;  PRIOR FILING DATE: 2003-03-20
;  NUMBER OF SEQ ID NOS: 64
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo spaiens
US-10-407-449-6

Alignment Scores:
Pred. No.:              487           Length:        25
Score:                  32.00         Matches:       5
Percent Similarity:     85.71%        Conservative:  1
Best Local Similarity:  71.43%        Mismatches:    1
Query Match:            78.05%        Indels:        0
DB:                     16            Gaps:          0

US-09-606-910E-3 (1-7) x US-10-407-449-6 (1-25)

Qy          1 ProAspThrArgProAlaPro 7
              ||||||:::||||||   |||
Db         24 CCCGACTCCCGCCCCCTTCCT 4

RESULT 2
US-10-660-897-6/c
; Sequence 6, Application US/10660897
; Publication No. US20040115706A1
; GENERAL INFORMATION:
;  APPLICANT: Jin, Cheng
;  APPLICANT:  Chung, Mary
;  APPLICANT:  Siddiqui-Jain, Adam
;  APPLICANT:  Whitten, Jeffrey
;  APPLICANT:  Farrell, Thomas
;  TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR IDENTIFYING
;  TITLE OF INVENTION:  QUADRUPLEX FORMING NUCLEIC ACIDS AND MODULATORS THEREOF
;  FILE REFERENCE: 532232000800
;  CURRENT APPLICATION NUMBER: US/10/660,897
;  CURRENT FILING DATE:  2003-09-11
;  PRIOR APPLICATION NUMBER: 60/410,475
;  PRIOR FILING DATE: 2002-09-12
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: quadruplex forming sequence
US-10-660-897-6

Alignment Scores:
Pred. No.:              487           Length:        25
Score:                  32.00         Matches:       5
Percent Similarity:     85.71%        Conservative:  1
Best Local Similarity:  71.43%        Mismatches:    1
Query Match:            78.05%        Indels:        0
DB:                     17            Gaps:          0

US-09-606-910E-3 (1-7) x US-10-660-897-6 (1-25)

Qy          1 ProAspThrArgProAlaPro 7
              ||||||:::||||||   |||
Db         24 CCCGACTCCCGCCCCCTTCCT 4

RESULT 3
US-10-273-180-19
; Sequence 19, Application US/10273180
; Publication No. US20030220476A1
; GENERAL INFORMATION:
;  APPLICANT: KOH, Gou Young
;  TITLE OF INVENTION: CHIMERIC COILED COIL MOLECULES
;  FILE REFERENCE: 10010-00001
;  CURRENT APPLICATION NUMBER: US/10/273,180
;  CURRENT FILING DATE:  2002-10-18
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: Primer for COMP
US-10-273-180-19

Alignment Scores:
Pred. No.:              568           Length:        30
Score:                  32.00         Matches:       5
Percent Similarity:     85.71%        Conservative:  1
Best Local Similarity:  71.43%        Mismatches:    1
Query Match:            78.05%        Indels:        0
DB:                     15            Gaps:          0

US-09-606-910E-3 (1-7) x US-10-273-180-19 (1-30)

Qy          1 ProAspThrArgProAlaPro 7
              ||||||   ||||||:::|||
Db          1 CCAGATCTTAGACCTAGCCCC 21

RESULT 4
US-09-894-633A-37
; Sequence 37, Application US/09894633A
; Patent No. US20020124285A1
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy
;  APPLICANT:  Dubois, Patrice
;  APPLICANT:  Malven, Marianne
;  APPLICANT:  Masucci, James
;  TITLE OF INVENTION: PLANT REGULATORY SEQUENCES FOR SELECTIVE CONTROL OF GENE EXPRESSI
;  FILE REFERENCE: 38-21(15856)B
;  CURRENT APPLICATION NUMBER: US/09/894,633A
;  CURRENT FILING DATE:  2001-06-28
;  PRIOR APPLICATION NUMBER: 60/214,357
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 09/894,633
;  PRIOR FILING DATE: 2000-06-28
;  NUMBER OF SEQ ID NOS: 111
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
;   LENGTH: 39
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(39)
;   OTHER INFORMATION: synthetic primer sequence
US-09-894-633A-37

Alignment Scores:
Pred. No.:              706           Length:        39
Score:                  32.00         Matches:       5
Percent Similarity:     85.71%        Conservative:  1
Best Local Similarity:  71.43%        Mismatches:    1
Query Match:            78.05%        Indels:        0
DB:                     9             Gaps:          0

US-09-606-910E-3 (1-7) x US-09-894-633A-37 (1-39)

Qy          1 ProAspThrArgProAlaPro 7
              ||||||   ||||||:::|||
Db          5 CCAGATCTGAGGCCCTCGCCT 25

RESULT 5
US-09-993-346-74
; Sequence 74, Application US/09993346
; Publication No. US20030124530A1
;   GENERAL INFORMATION:
;        APPLICANT: Edwards, Cynthia A.
;                   Cantor, Charles R.
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         December 29, 2004, 15:46:38 ; Search time 69 Seconds
                                          (without alignments)
                                          72.109 Million cell updates/sec

Title:          US-09-606-910E-3
Perfect score:  41
Sequence:       1 PDTRPAP 7

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters:      905748

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09606910/runat_29122004_131141_4709/app_query.fasta_1.199
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=50nt.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=50
-USER=US09606910_@CGN_1_1_69_@runat_29122004_131141_4709 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                2:  /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                3:  /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                6:  /cgn2_6/ptodata/1/ina/backfiles1.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                        Description
 -------------------------------------------------------------------------
     1     32    78.0     47  1  US-08-171-389-74           Sequence 74, Appl
     2     32    78.0     47  1  US-08-123-936-74           Sequence 74, Appl
     3     32    78.0     47  2  US-08-475-228A-74          Sequence 74, Appl
     4     32    78.0     47  3  US-08-482-080A-74          Sequence 74, Appl
     5     32    78.0     47  3  US-09-354-947-74           Sequence 74, Appl
     6     32    78.0     47  5  PCT-US93-12388-74          Sequence 74, Appl
     7     32    78.0     50  1  US-08-171-389-95           Sequence 95, Appl
     8     32    78.0     50  1  US-08-123-936-95           Sequence 95, Appl
     9     32    78.0     50  2  US-08-475-228A-95          Sequence 95, Appl
    10     32    78.0     50  3  US-08-482-080A-95          Sequence 95, Appl
    11     32    78.0     50  3  US-09-354-947-95           Sequence 95, Appl
    12     32    78.0     50  5  PCT-US93-12388-95          Sequence 95, Appl
```

```
     13      31    75.6     40  1  US-08-030-731A-25          Sequence 25, Appl
c    14      31    75.6     40  1  US-08-030-731A-26          Sequence 26, Appl
     15      31    75.6     46  1  US-08-030-731A-28          Sequence 28, Appl
c    16      31    75.6     46  1  US-08-030-731A-29          Sequence 29, Appl
     17      31    75.6     46  1  US-08-030-731A-31          Sequence 31, Appl
c    18      31    75.6     46  1  US-08-030-731A-32          Sequence 32, Appl
c    19      30    73.2     20  3  US-08-974-549A-458         Sequence 458, App
c    20      30    73.2     20  4  US-08-912-951-225          Sequence 225, App
c    21      30    73.2     20  4  US-09-402-181B-458         Sequence 458, App
c    22      30    73.2     20  4  US-09-721-456-458          Sequence 458, App
     23      30    73.2     25  3  US-08-974-549A-441         Sequence 441, App
     24      30    73.2     25  4  US-08-912-951-208          Sequence 208, App
     25      30    73.2     25  4  US-09-402-181B-441         Sequence 441, App
     26      30    73.2     25  4  US-09-721-456-441          Sequence 441, App
     27      30    73.2     26  4  US-09-066-357-9            Sequence 9, Appli
c    28      30    73.2     31  4  US-09-733-042-37           Sequence 37, Appl
c    29      30    73.2     42  1  US-08-530-492-46           Sequence 46, Appl
c    30      30    73.2     42  3  US-08-906-517-46           Sequence 46, Appl
c    31      30    73.2     45  4  US-09-868-758-15           Sequence 15, Appl
c    32      30    73.2     45  4  US-09-868-758-17           Sequence 17, Appl
     33      29    70.7     25  4  US-08-714-741-22           Sequence 22, Appl
     34      29    70.7     26  1  US-08-072-070-5            Sequence 5, Appli
     35      29    70.7     26  1  US-08-469-434-5            Sequence 5, Appli
     36      29    70.7     26  1  US-08-214-222-5            Sequence 5, Appli
     37      29    70.7     26  2  US-08-467-852A-6           Sequence 6, Appli
     38      29    70.7     26  2  US-08-468-718-5            Sequence 5, Appli
     39      29    70.7     26  2  US-08-247-491A-6           Sequence 6, Appli
     40      29    70.7     26  2  US-08-319-795-3            Sequence 3, Appli
     41      29    70.7     26  2  US-08-468-985-3            Sequence 3, Appli
     42      29    70.7     26  3  US-08-312-949-5            Sequence 5, Appli
     43      29    70.7     26  4  US-08-714-741-1            Sequence 1, Appli
     44      29    70.7     26  4  US-08-714-741-19           Sequence 19, Appl
     45      29    70.7     26  4  US-08-529-055-1            Sequence 1, Appli

                                  ALIGNMENTS

RESULT 1
US-08-171-389-74
; Sequence 74, Application US/08171389
; Patent No. 5578444
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/171,389
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175/G19P3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  74:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  47 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human complement C3 gene
US-08-171-389-74

Alignment Scores:
Pred. No.:              281            Length:         47
Score:                  32.00          Matches:        5
Percent Similarity:     85.71%         Conservative:   1
Best Local Similarity:  71.43%         Mismatches:     1
Query Match:            78.05%         Indels:         0
DB:                     1              Gaps:           0

US-09-606-910E-3 (1-7) x US-08-171-389-74 (1-47)

Qy          1 ProAspThrArgProAlaPro 7
              ||||||   :::|||||||||
Db         17 CCAGATAAAAAGCCAGCTCCA 37

RESULT 2
US-08-123-936-74
; Sequence 74, Application US/08123936
; Patent No. 5726014
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    TITLE OF INVENTION:  Screening Assay for the Detection of
;    TITLE OF INVENTION:  DNA-Binding Molecules
;    NUMBER OF SEQUENCES:  640
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/123,936
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
```

```
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0075.32/G19P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  74:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  47 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human complement C3 gene
US-08-123-936-74

Alignment Scores:
Pred. No.:              281            Length:         47
Score:                  32.00          Matches:        5
Percent Similarity:     85.71%         Conservative:   1
Best Local Similarity:  71.43%         Mismatches:     1
Query Match:            78.05%         Indels:         0
DB:                     1              Gaps:           0

US-09-606-910E-3 (1-7) x US-08-123-936-74 (1-47)

Qy          1 ProAspThrArgProAlaPro 7
              ||||||   :::|||||||||
Db         17 CCAGATAAAAAGCCAGCTCCA 37

RESULT 3
US-08-475-228A-74
; Sequence 74, Application US/08475228A
; Patent No. 5869241
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,228A
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:          December 29, 2004, 14:55:26 ; Search time 330 Seconds
                                             (without alignments)
                                             111.351 Million cell updates/sec

Title:           US-09-606-910E-3
Perfect score:   41
Sequence:        1 PDTRPAP 7

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters:      3992700

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09606910/runat_29122004_131139_4673/app_query.fasta_1.199
-DB=N_Geneseq_23Sep04 -QFMT=fastap -SUFFIX=50nt.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=50
-USER=US09606910_@CGN_1_1_470_@runat_29122004_131139_4673 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       N_Geneseq_23Sep04:*
                 1:  geneseqn1980s:*
                 2:  geneseqn1990s:*
                 3:  geneseqn2000s:*
                 4:  geneseqn2001as:*
                 5:  geneseqn2001bs:*
                 6:  geneseqn2002as:*
                 7:  geneseqn2002bs:*
                 8:  geneseqn2003as:*
                 9:  geneseqn2003bs:*
                 10:  geneseqn2003cs:*
                 11:  geneseqn2003ds:*
                 12:  geneseqn2004s:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                    SUMMARIES
                   %
Result           Query
  No.    Score   Match Length DB  ID                     Description
 ---------------------------------------------------------------------------
     1     41    100.0     48  4  AAS01380               Aas01380 Human muc
     2     34     82.9     21  2  AAQ04708               Aaq04708 Sequence
     3     34     82.9     39  2  AAT90104               Aat90104 MUC1 modi
c    4     33     80.5     41  4  AAH46801               Aah46801 Human qui
c    5     32     78.0     25 10   ADE86161               Ade86161 BCL2 gene
     6     32     78.0     30 12   ADI23653               Adi23653 Human car
```

```
     7       32    78.0     37  12  ADO80343             Ado80343 Mouse pho
     8       32    78.0     39   6  ABK27844             Abk27844 Corn male
c    9       32    78.0     41   4  AAH46800             Aah46800 Human qui
c   10       32    78.0     41   6  ABZ50187             Abz50187 Human NDU
c   11       32    78.0     41   6  ABZ44175             Abz44175 Human NDU
    12       32    78.0     47   2  AAQ69324             Aaq69324 Human com
    13       32    78.0     47   2  AAT63786             Aat63786 Human com
    14       32    78.0     47   2  AAX17074             Aax17074 Test sequ
    15       32    78.0     47   6  ABK82565             Abk82565 DNA bindi
    16       32    78.0     47  12   ADE80104             Ade80104 Duplex ol
    17       32    78.0     50   2  AAQ69345             Aaq69345 Human C3
    18       32    78.0     50   2  AAT63807             Aat63807 Human C3
    19       32    78.0     50   2  AAX17095             Aax17095 Test sequ
    20       32    78.0     50   6  ABK82586             Abk82586 DNA bindi
    21       32    78.0     50  12   ADE80125             Ade80125 Duplex ol
    22       31    75.6     25   9  ACK10224             Ack10224 Human mic
    23       31    75.6     25   9  ACK10225             Ack10225 Human mic
c   24       31    75.6     25   9  ACK05331             Ack05331 Human mic
c   25       31    75.6     25   9  ACI62281             Aci62281 Human mic
c   26       31    75.6     25   9  ACI47210             Aci47210 Human mic
c   27       31    75.6     27   2  AAT76379             Aat76379 Human GM-
c   28       31    75.6     27   2  AAX54178             Aax54178 Human GM-
c   29       31    75.6     27   3  AAA33622             Aaa33622 Low adeno
c   30       31    75.6     27   3  AAF19744             Aaf19744 Human GM-
c   31       31    75.6     27  10   ABZ95438             Abz95438 Human GM-
c   32       31    75.6     27  11   ABD19521             Abd19521 Human GM-
    33       31    75.6     40   2  AAQ14250             Aaq14250 Proinsuli
    34       31    75.6     41   6  ABZ48607             Abz48607 Human N-m
    35       31    75.6     41   6  ABZ43312             Abz43312 Human N-m
    36       31    75.6     46   2  AAQ14252             Aaq14252 Proinsuli
    37       31    75.6     46   2  AAQ14251             Aaq14251 Proinsuli
    38       31    75.6     47   3  AAZ67256             Aaz67256 Human map
c   39       30    73.2     17   8  ABZ62078             Abz62078 Human H-R
    40       30    73.2     19   2  AAT90102             Aat90102 Human epi
c   41       30    73.2     20   2  AAV30681             Aav30681 Telomeras
    42       30    73.2     20  12   ADP76557             Adp76557 Chimeric
    43       30    73.2     20  12   ADP76657             Adp76657 Chimeric
    44       30    73.2     20  12   ADP76492             Adp76492 Chimeric
    45       30    73.2     25   2  AAV30663             Aav30663 Telomeras

                              ALIGNMENTS

RESULT 1
AAS01380
ID   AAS01380 standard; DNA; 48 BP.
XX
AC   AAS01380;
XX
DT   04-JUL-2001  (first entry)
XX
DE   Human mucin peptide MUC1(16) encoding DNA sequence.
XX
KW   Human; polymorphic epithelial cell mucin; PEM; chimeric virus particle;
KW   CVP; plant virus coat protein; comovirus; CPMV; cowpea mosaic virus;
KW   SBMV; Southern bean mosaic virus; LTSV; red clover necrotic mosaic virus;
KW   RCNMV; MUC1; tumour; cancer vaccine; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..48
FT                   /*tag=  a
FT                   /partial
FT                   /product= "MUC1(16) peptide"
FT                   /note= "This sequence lacks both start and stop codons"
XX
PN   WO200118199-A1.
XX
PD   15-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-GB003500.
XX
```

```
PR   09-SEP-1999;   99GB-00021337.
XX
PA   (DOWC ) DOW CHEM CO.
XX
PI   Bendig MM,  Jones TD,  Longstaff M,  Hellendoorn K;
XX
DR   WPI; 2001-244570/25.
DR   P-PSDB; AAU00488.
XX
PT   New chimeric plant virus particles with an immunogenically active peptide
PT   of a tumor-associated mucin, useful as a vaccine or for the manufacture
PT   of a vaccine for treating and/or preventing tumors and/or cancer.
XX
PS   Example 12; Fig 8B; 63pp; English.
XX
CC   The present sequence encoding for human mucin peptide MUC1(16) is
CC   inserted into an immunogenic site encoding for the coat protein of
CC   several different plant viruses to create a novel chimeric virus particle
CC   (CVP). Constructs containing the MUC1(16) DNA sequence at different
CC   positions within several plant virus coat protein insertion sequences are
CC   created. MUC1(16) peptide is 1 of 7 polymorphic epithelial cell mucin
CC   (PEM) peptides (AAU00483-AAU00489) which can be used to create CVPs. The
CC   mucin peptide epitopes are preferably inserted into the coat protein of a
CC   plant virus such as the comovirus CPMV (cowpea mosaic virus). The ability
CC   of the chimeric virus particle CPMV-MUC1(16) to elicit antibodies, which
CC   can cause regression of tumours expressing the MUC1 protein, is
CC   demonstrated in a mouse tumour model. Other examples of CVPs include the
CC   insertion of the mucin MUC1(16) peptide into the coat protein of other
CC   plant viruses e.g. SBMV (Southern bean mosaic virus; AAU00490), LTSV
CC   (AAU00491) and RCNMV (red clover necrotic mosaic virus; AAU00492). The
CC   CVP is useful as or as part of a vaccine particularly for treating and
CC   preventing tumours and cancer. The CVP provides advantages over prior art
CC   antigen-presenting means since conventional live animal virus vectors can
CC   be avoided; as can the need for separate mucin peptide synthesis and
CC   chemical-coupling to a conventional carrier. Also, the CVP is shown to
CC   induce good mucosal immunity, and does not require the addition of
CC   exogenous adjuvants to induce a strong immune response
XX
SQ   Sequence 48 BP; 6 A; 14 C; 10 G; 18 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              60.9            Length:         48
Score:                  41.00           Matches:        7
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     4               Gaps:           0

US-09-606-910E-3 (1-7) x AAS01380 (1-48)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         16 CCTGATACTAGACCTGCTCCT 36

RESULT 2
AAQ04708
ID   AAQ04708 standard; DNA; 21 BP.
XX
AC   AAQ04708;
XX
DT   25-MAR-2003  (revised)
DT   10-OCT-1990  (first entry)
XX
DE   Sequence encoding epitopic fragment of polymorphic epithelial mucin
DE   (PEM).
XX
KW   Breast cancer; cancer; polymorphic epithelial mucin; mucin; PEM.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   repeat_region   1..3
FT                   /*tag=  a
```

```
FT   repeat_region   19..21
FT                   /*tag=  a
XX
PN   WO9005142-A.
XX
PD   17-MAY-1990.
XX
PF   10-NOV-1988;   88GB-00026362.
XX
PR   10-NOV-1988;   88GB-00026362.
PR   05-APR-1989;   89GB-00007660.
XX
PA   (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.
XX
PI   Taylorpapa J,  Gendler S,  Burchell J;
XX
DR   WPI; 1990-178797/23.
XX
PT   Polypeptide(s) having sequence of antigenic epitope on PEM - (polymorphic
PT   epithelial mucin), useful for producing antibodies for diagnosis and
PT   therapy of breast cancer.
XX
PS   Disclosure; Page ?; -pp; English.
XX
CC   The protein fragment encoded by the sequence lying between the two
CC   repeated regions (pdtrp) is an epitope in PEM which is expressed at high
CC   levels in sufferers from breast cancer, and is thus useful in generating
CC   antibodies for diagnosis and therapy. (Updated on 25-MAR-2003 to correct
CC   PR field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-
CC   MAR-2003 to correct PI field.)
XX
SQ   Sequence 21 BP; 3 A; 11 C; 7 G; 0 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              366             Length:         21
Score:                  34.00           Matches:        6
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            82.93%          Indels:         0
DB:                     2               Gaps:           0

US-09-606-910E-3 (1-7) x AAQ04708 (1-21)

Qy          1 ProAspThrArgProAla 6
              ||||||||||||||||||
Db          4 CCGGACACCAGGCCGGCC 21

RESULT 3
AAT90104
ID   AAT90104 standard; DNA; 39 BP.
XX
AC   AAT90104;
XX
DT   02-APR-1998  (first entry)
XX
DE   MUC1 modified epitope amplifying primer PRMUCIM-5'.
XX
KW   Hepatitis B virus core antigen; HBcAg; MUC1; tumour antigen epitope;
KW   chimeric; Ha-ras oncogene; mucin; tumour; human; PCR primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9735008-A1.
XX
PD   25-SEP-1997.
XX
PF   21-MAR-1997;   97WO-US004656.
XX
PR   21-MAR-1996;   96US-0013839P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
```

```
                    GenCore version 5.1.6
             Copyright (c) 1993 - 2004  Compugen Ltd.


OM protein - nucleic search, using frame_plus_p2n model

Run on:         December 29, 2004, 15:07:52 ; Search time 2542 Seconds
                                              (without alignments)
                                              130.223 Million cell updates/sec

Title:          US-09-606-910E-3
Perfect score:  41
Sequence:       1 PDTRPAP 7

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters:      1785872

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09606910/runat_29122004_131140_4683/app_query.fasta_1.199
-DB=GenEmbl -QFMT=fastap -SUFFIX=50nt.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=50
-USER=US09606910_@CGN_1_1_3731_@runat_29122004_131140_4683 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES

                   %
Result           Query
  No.    Score   Match Length DB  ID                            Description
 ------------------------------------------------------------------------------
c  1        41   100.0     35  6  A21776                        A21776 oligonucleo
c  2        41   100.0     35  6  AX192399                      AX192399 Sequence
c  3        41   100.0     35  6  BD000574                      BD000574 Human pol
   4        34    82.9     21  6  A03527                        A03527 Markush syn
c  5        34    82.9     30  6  AX192395                      AX192395 Sequence
c  6        34    82.9     30  6  BD000570                      BD000570 Human pol
   7        32    78.0     37  6  CQ788435                      CQ788435 Sequence
   8        32    78.0     39  6  AX356243                      AX356243 Sequence
   9        32    78.0     47  6  AR032462                      AR032462 Sequence
  10        32    78.0     47  6  I29202                        I29202 Sequence 74
  11        32    78.0     47  6  I90876                        I90876 Sequence 74
  12        32    78.0     47  6  AR209126                      AR209126 Sequence
  13        32    78.0     50  6  AR032483                      AR032483 Sequence
  14        32    78.0     50  6  I29223                        I29223 Sequence 95
  15        32    78.0     50  6  I90897                        I90897 Sequence 95
  16        32    78.0     50  6  AR209147                      AR209147 Sequence
  17        31    75.6     40  6  I12504                        I12504 Sequence 25
c 18        31    75.6     40  6  I12505                        I12505 Sequence 26
  19        31    75.6     41  6  AX513898                      AX513898 Sequence
c 20        31    75.6     41  6  AX514053                      AX514053 Sequence
c 21        31    75.6     41  6  AX516139                      AX516139 Sequence
c 22        31    75.6     41  6  AX517544                      AX517544 Sequence
  23        31    75.6     41  6  AX519192                      AX519192 Sequence
c 24        31    75.6     41  6  AX520522                      AX520522 Sequence
  25        31    75.6     46  6  I12506                        I12506 Sequence 28
c 26        31    75.6     46  6  I12507                        I12507 Sequence 29
  27        31    75.6     46  6  I12508                        I12508 Sequence 31
c 28        31    75.6     46  6  I12509                        I12509 Sequence 32
c 29        30    73.2     20  6  E36911                        E36911 Human telom
c 30        30    73.2     20  6  AR243432                      AR243432 Sequence
c 31        30    73.2     20  6  AR390588                      AR390588 Sequence
c 32        30    73.2     20  6  AR393202                      AR393202 Sequence
c 33        30    73.2     20  6  AX810493                      AX810493 Sequence
c 34        30    73.2     20  6  BD011162                      BD011162 Human tel
  35        30    73.2     25  6  E36894                        E36894 Human telom
  36        30    73.2     25  6  AR243415                      AR243415 Sequence
  37        30    73.2     25  6  AR390571                      AR390571 Sequence
  38        30    73.2     25  6  AR393185                      AR393185 Sequence
  39        30    73.2     25  6  AX810476                      AX810476 Sequence
  40        30    73.2     25  6  BD011145                      BD011145 Human tel
c 41        30    73.2     26  6  CQ753464                      CQ753464 Sequence
  42        30    73.2     26  6  AR374981                      AR374981 Sequence
c 43        30    73.2     27  6  BD225170                      BD225170 Medicinal
c 44        30    73.2     28  6  CQ753465                      CQ753465 Sequence
  45        30    73.2     29  6  BD105752                      BD105752 Ciliary n


                                ALIGNMENTS


RESULT 1
A21776/c
LOCUS       A21776                    35 bp    DNA     linear   PAT 21-JUL-1994
DEFINITION  oligonucleotide.
ACCESSION   A21776
VERSION     A21776.1  GI:583655
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 35)
  AUTHORS   .
  TITLE     MUCIN NUCLEOTIDES
  JOURNAL   Patent: WO 9109867-A 13 11-JUL-1991;
FEATURES             Location/Qualifiers
     source          1. .35
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

Alignment Scores:
Pred. No.:              43.5          Length:         35
Score:                  41.00         Matches:        7
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     6             Gaps:           0
```

```
US-09-606-910E-3 (1-7) x A21776 (1-35)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         35 CCGGACACCAGGCCGGCCCCG 15

RESULT 2
AX192399/c
LOCUS       AX192399                  35 bp    DNA     linear   PAT 15-AUG-2001
DEFINITION  Sequence 5 from Patent EP1103623.
ACCESSION   AX192399
VERSION     AX192399.1  GI:15210366
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Taylor-Papadimitriou,J., Burchell,J. and Gendler,S.
  TITLE     Human mucin core protein: nucleic acid probes, peptide fragments
            and antibodies thereto, and uses thereof in diagnostic and
            therapeutic methods
  JOURNAL   Patent: EP 1103623-A 5 30-MAY-2001;
            IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1..35
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              43.5          Length:        35
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     6             Gaps:          0

US-09-606-910E-3 (1-7) x AX192399 (1-35)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         21 CCGGACACCAGGCCGGCCCCG 1

RESULT 3
BD000574/c
LOCUS       BD000574                  35 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Human polymorphic epithelial mucin core protein and nucleic acid
            encoding  the protein.
ACCESSION   BD000574
VERSION     BD000574.1  GI:18623687
KEYWORDS    JP 2000333675-A/5.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 35)
  AUTHORS   Papadimitrov,J.T., Jendora,S. and Bachieru,J.
  TITLE     Human polymorphic epithelial mucin core protein and nucleic acid
            encoding the protein
  JOURNAL   Patent: JP 2000333675-A 5 05-DEC-2000;
            IMPERIAL CANCER RESEARCH TECHNOLOGY LTD
COMMENT     OS   Homo sapiens (human)
            PN   JP 2000333675-A/5
            PD   05-DEC-2000
            PF   26-APR-2000 JP 2000125724
            PR   07-JAN-1987 GB       8700269,07-JAN-1987 GB       8700279 PR
            22-APR-1987 US       041306,09-NOV-1987 GB       8726172 PI   JOYCE
            TAYLOR PAPADIMITROV,SANDRA JENDORA,JOY BACHIERU PC
            C12N15/02,A61K38/00,A61K39/395,A61K39/395,A61K49/00,A61P35/00, PC
            C07K14/47,
```

```
            PC   C07K16/44,C12N5/10,C12P21/08//(C12N15/02,C12R1:91),(C12N5/10,
            PC   C12R1:91),
            PC   C12N15/00,A61K37/02,C12N5/00,(C12N15/00,C12R1:91),(C12N5/00,
            PC   C12R1:91)
            CC
            FH   Key             Location/Qualifiers
            FT   source          1..35
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1..35
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              43.5          Length:        35
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     6             Gaps:          0

US-09-606-910E-3 (1-7) x BD000574 (1-35)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         21 CCGGACACCAGGCCGGCCCCG 1

RESULT 4
A03527
LOCUS       A03527                    21 bp    DNA     linear   PAT 17-MAR-1993
DEFINITION  Markush synthetic DNA sequence.
ACCESSION   A03527
VERSION     A03527.1  GI:344686
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   .
  TITLE     POLYPEPTIDES
  JOURNAL   Patent: WO 9005142-A 9 17-MAY-1990;
FEATURES             Location/Qualifiers
     source          1..21
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
     repeat_region   1..3
                     /note="region may not be present, occur once or several
                     times"
     repeat_region   19..21
                     /note="region may not be present, occur once or several
                     times"
ORIGIN

Alignment Scores:
Pred. No.:              431           Length:        21
Score:                  34.00         Matches:       6
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            82.93%        Indels:        0
DB:                     6             Gaps:          0

US-09-606-910E-3 (1-7) x A03527 (1-21)

Qy          1 ProAspThrArgProAla 6
              ||||||||||||||||||
Db          4 CCGGACACCAGGCCGGCC 21

RESULT 5
AX192395/c
LOCUS       AX192395                  30 bp    DNA     linear   PAT 15-AUG-2001
```

```
                     GenCore version 5.1.6
              Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         December 29, 2004, 13:16:09 ; Search time 89 Seconds
                                          (without alignments)
                                          45.254 Million cell updates/sec

Title:          US-09-606-910E-3
Perfect score:  41
Sequence:       1 PDTRPAP 7

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters:      79092

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       UniProt_02:*
                1:  uniprot_sprot:*
                2:  uniprot_trembl:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                   SUMMARIES
                 %
Result           Query
  No.   Score   Match Length DB  ID                      Description
----------------------------------------------------------------------------
    1      41   100.0     20  2  Q9UMI8                  Q9umi8 homo sapien
    2      29    70.7     15  1  MK1_PALPR               P80408 palomena pr
    3      29    70.7     15  1  MK2A_PALPR              P80409 palomena pr
    4      29    70.7     16  1  MK2B_PALPR              P80410 palomena pr
    5      29    70.7     16  1  MK3_PALPR               P80411 palomena pr
    6      29    70.7     33  2  O77492                  O77492 nycticebus
    7      29    70.7     43  2  Q6CI56                  Q6ci56 yarrowia li
    8      28    68.3     35  2  Q8FZI5                  Q8fzi5 brucella su
    9      27    65.9     31  2  Q981P2                  Q981p2 rhizobium l
   10      27    65.9     31  2  P89087                  P89087 hepatitis d
   11      27    65.9     31  2  P89089                  P89089 hepatitis d
   12      27    65.9     43  2  Q14051                  Q14051 homo sapien
   13      27    65.9     47  2  Q8NDV6                  Q8ndv6 homo sapien
   14      27    65.9     48  1  TXA1_STOHE              P19651 stoichactis
   15      26    63.4     20  2  Q8V9H3                  Q8v9h3 chicken ane
   16      26    63.4     26  1  LCLP_HUMAN              P34168 homo sapien
   17      26    63.4     34  2  Q9QVB3                  Q9qvb3 rattus sp.
   18      26    63.4     44  2  O35081                  O35081 mus musculu
   19      26    63.4     47  2  Q747Q0                  Q747q0 geobacter s
   20      26    63.4     47  2  AAR36606                Aar36606 geobacter
   21      25    61.0     16  2  Q9UQS4                  Q9uqs4 homo sapien
   22      25    61.0     31  2  Q6XGD1                  Q6xgd1 leifsonia x
   23      25    61.0     31  2  AAP69566                Aap69566 leifsonia
   24      25    61.0     35  2  Q97SL6                  Q97sl6 streptococc
   25      25    61.0     37  2  Q77132                  Q77132 human immun
   26      25    61.0     41  2  Q9AD27                  Q9ad27 streptomyce
   27      25    61.0     42  2  Q16469                  Q16469 homo sapien
   28      25    61.0     46  2  Q7M1G3                  Q7m1g3 lolium mult
   29      25    61.0     47  2  P70886                  P70886 butyrivibri
   30      25    61.0     50  2  Q9TBA8                  Q9tba8 mergellus a
   31      25    61.0     50  2  Q8U5M1                  Q8u5m1 agrobacteri
   32      25    61.0     50  2  Q8YCK3                  Q8yck3 brucella me
   33    24.5    59.8     46  2  Q7R732                  Q7r732 plasmodium
   34      24    58.5     16  2  Q9TRR1                  Q9trr1 oryctolagus
   35      24    58.5     17  2  Q9UDD6                  Q9udd6 homo sapien
   36      24    58.5     19  2  Q8TB44                  Q8tb44 homo sapien
   37      24    58.5     22  1  WAG1_TRIWA              P24335 trimeresuru
   38      24    58.5     22  1  WAG2_TRIWA              P58930 trimeresuru
   39      24    58.5     22  2  Q6JDI1                  Q6jdi1 canis famil
   40      24    58.5     23  2  Q9TRR0                  Q9trr0 oryctolagus
   41      24    58.5     25  2  Q69148                  Q69148 human herpe
   42      24    58.5     25  2  Q69474                  Q69474 human herpe
   43      24    58.5     28  2  Q9HDB7                  Q9hdb7 homo sapien
   44      24    58.5     33  2  Q9RMS9                  Q9rms9 pantoea agg
   45      24    58.5     34  2  Q7RXS6                  Q7rxs6 neurospora

                                  ALIGNMENTS

RESULT 1
Q9UMI8
ID   Q9UMI8      PRELIMINARY;      PRT;    20 AA.
AC   Q9UMI8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Mucin (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89235154; PubMed=2715633;
RA   Xing P.X., Tjandra J.J., Reynolds K., McLaughlin P.J., Purcell D.F.J.,
RA   McKenzie I.F.C.;
RT   "Reactivity of anti-human milk fat globule antibodies with synthetic
RT   peptides.";
RL   J. Immunol. 142:3503-3509(1989).
DR   EMBL; M26316; AAA36336.1; -.
DR   PIR; S10571; S10571.
FT   NON_TER       1      1
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  1887 MW;  5B3473EAEBAFAD87 CRC64;

  Query Match            100.0%;  Score 41;  DB 2;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 1.6;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          1 PDTRPAP 7


RESULT 2
MK1_PALPR
ID   MK1_PALPR      STANDARD;      PRT;    15 AA.
AC   P80408;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Metalnikowin I.
OS   Palomena prasina (Green shield bug).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;
OC   Panheteroptera; Pentatomomorpha; Pentatomoidea; Pentatomidae;
OC   Pentatominae; Palomena.
OX   NCBI_TaxID=55431;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Hemolymph;
RA   Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;
RT   "The inducible antibacterial peptides of the hemipteran insect
```

```
RT   Palomena prasina: identification of a unique family of proline-rich
RT   peptides and of a novel insect defensin.";
RL   J. Insect Physiol. 42:81-89(1996).
CC   -!- FUNCTION: Antibacterial peptide active against Gram-negative
CC       bacteria.
CC   -!- INDUCTION: By bacterial infection.
KW   Antibiotic; Direct protein sequencing; Insect immunity.
SQ   SEQUENCE   15 AA;  1838 MW;  21407E663CE46299 CRC64;

  Query Match             70.7%;  Score 29;  DB 1;  Length 15;
  Best Local Similarity   71.4%;  Pred. No. 1.7e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              || || |
Db          4 PDYRPRP 10


RESULT 3
MK2A_PALPR
ID   MK2A_PALPR     STANDARD;      PRT;    15 AA.
AC   P80409;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Metalnikowin IIA.
OS   Palomena prasina (Green shield bug).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;
OC   Panheteroptera; Pentatomomorpha; Pentatomoidea; Pentatomidae;
OC   Pentatominae; Palomena.
OX   NCBI_TaxID=55431;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Hemolymph;
RA   Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;
RT   "The inducible antibacterial peptides of the hemipteran insect
RT   Palomena prasina: identification of a unique family of proline-rich
RT   peptides and of a novel insect defensin.";
RL   J. Insect Physiol. 42:81-89(1996).
CC   -!- FUNCTION: Antibacterial peptide active against Gram-negative
CC       bacteria.
CC   -!- INDUCTION: By bacterial infection.
KW   Antibiotic; Direct protein sequencing; Insect immunity.
SQ   SEQUENCE   15 AA;  1893 MW;  23835D063B946299 CRC64;

  Query Match             70.7%;  Score 29;  DB 1;  Length 15;
  Best Local Similarity   71.4%;  Pred. No. 1.7e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              || || |
Db          4 PDYRPRP 10


RESULT 4
MK2B_PALPR
ID   MK2B_PALPR     STANDARD;      PRT;    16 AA.
AC   P80410;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Metalnikowin IIB.
OS   Palomena prasina (Green shield bug).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;
OC   Panheteroptera; Pentatomomorpha; Pentatomoidea; Pentatomidae;
OC   Pentatominae; Palomena.
OX   NCBI_TaxID=55431;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Hemolymph;
RA   Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;
```

```
RT   "The inducible antibacterial peptides of the hemipteran insect
RT   Palomena prasina: identification of a unique family of proline-rich
RT   peptides and of a novel insect defensin.";
RL   J. Insect Physiol. 42:81-89(1996).
CC   -!- FUNCTION: Antibacterial peptide active against Gram-negative
CC       bacteria.
CC   -!- INDUCTION: By bacterial infection.
KW   Antibiotic; Direct protein sequencing; Insect immunity.
SQ   SEQUENCE   16 AA;  2040 MW;  AF21407D063B9462 CRC64;

  Query Match             70.7%;  Score 29;  DB 1;  Length 16;
  Best Local Similarity   71.4%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              || || |
Db          4 PDYRPRP 10


RESULT 5
MK3_PALPR
ID   MK3_PALPR      STANDARD;      PRT;    16 AA.
AC   P80411;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Metalnikowin III.
OS   Palomena prasina (Green shield bug).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;
OC   Panheteroptera; Pentatomomorpha; Pentatomoidea; Pentatomidae;
OC   Pentatominae; Palomena.
OX   NCBI_TaxID=55431;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Hemolymph;
RA   Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;
RT   "The inducible antibacterial peptides of the hemipteran insect
RT   Palomena prasina: identification of a unique family of proline-rich
RT   peptides and of a novel insect defensin.";
RL   J. Insect Physiol. 42:81-89(1996).
CC   -!- FUNCTION: Antibacterial peptide active against Gram-negative
CC       bacteria.
CC   -!- INDUCTION: By bacterial infection.
KW   Antibiotic; Direct protein sequencing; Insect immunity.
SQ   SEQUENCE   16 AA;  2024 MW;  A9E3835D063B9462 CRC64;

  Query Match             70.7%;  Score 29;  DB 1;  Length 16;
  Best Local Similarity   71.4%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              || || |
Db          4 PDYRPRP 10


RESULT 6
O77492
ID   O77492       PRELIMINARY;      PRT;    33 AA.
AC   O77492;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   D4 dopamine receptor (D4DR) (Fragment).
OS   Nycticebus coucang (Slow loris).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Strepsirhini; Loridae; Nycticebus.
OX   NCBI_TaxID=9470;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Inoue-Murayama M., Takenaka O., Murayama Y.;
RT   "Origin and divergence of tandem repeats of primate D4 dopamine
RT   receptor genes.";
```

```
                         GenCore version 5.1.6
               Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          December 29, 2004, 13:16:34 ; Search time 21 Seconds
                                         (without alignments)
                                         32.072 Million cell updates/sec

Title:           US-09-606-910E-3
Perfect score:   41
Sequence:        1 PDTRPAP 7

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:      11837

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        PIR_79:*
                 1:  pir1:*
                 2:  pir2:*
                 3:  pir3:*
                 4:  pir4:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                %
Result          Query
   No.   Score  Match Length DB  ID                 Description
-------------------------------------------------------------------------
     1      30   73.2     30  2  T30514             hypothetical prote
     2      28   68.3     25  2  S44201             HLA-DRB1 exon2 pro
     3      27   65.9     26  2  B60745             major outer membra
     4      27   65.9     48  1  A30114             long neurotoxin I
     5      26   63.4     26  2  S38295             serine proteinase
     6      25   61.0     35  2  C95037             hypothetical prote
     7      25   61.0     42  2  I70082             glycoprotein Ib al
     8      25   61.0     46  2  S07073             arabinogalactan pr
     9      25   61.0     47  2  T44626             hypothetical prote
    10      25   61.0     49  2  B24696             tissue kallikrein
    11      25   61.0     50  2  AD3575             hypothetical prote
    12      25   61.0     50  2  C97409             hypothetical prote
    13      24   58.5     13  2  S09716             2S albumin large c
    14      24   58.5     15  2  S67975             apolipoprotein Cb2
    15      24   58.5     17  2  A42920             fatty acid ethyl e
    16      24   58.5     19  2  B28457             proteoglycan II, b
    17      24   58.5     22  2  S67974             apolipoprotein Cb1
    18      24   58.5     23  2  B43836             fibronectin - rabb
    19      24   58.5     24  2  B44008             lethal peptide I -
    20      24   58.5     24  2  A44008             lethal peptide II
    21      24   58.5     25  2  I40692             cenA protein (IgA1
    22      24   58.5     34  2  A36626             osteogenic protein
    23      24   58.5     34  2  S02382             probable membrane
    24      24   58.5     36  2  C82475             hypothetical prote
    25      24   58.5     37  2  AG0540             hypothetical prote
    26      24   58.5     41  2  C87603             hypothetical prote
    27      24   58.5     41  2  A42064             lactam utilization
    28      24   58.5     45  2  S24713             Ig alpha chain, tr
    29      23   56.1     31  2  B18955             class II histocomp
```

```
    30      23   56.1     31  2  A25629             cytochrome-c oxida
    31      23   56.1     34  2  A19197             class II histocomp
    32      23   56.1     36  2  C37473             type I topoisomera
    33      23   56.1     36  2  T22263             hypothetical prote
    34      23   56.1     46  2  T28799             hypothetical prote
    35      23   56.1     47  2  JN0424             neurotoxin RTX-V -
    36      23   56.1     47  2  A40703             androgen-binding p
    37      23   56.1     48  1  TZAZR3             toxin RpIII - sea
    38      23   56.1     48  2  JN0425             neurotoxin RTX-IV
    39      23   56.1     49  2  S25829             serpin III - horse
    40      23   56.1     50  2  I80356             MHC HLA-DR-beta ch
    41      22   53.7     19  2  S12268             Qa-2 antigen - mou
    42      22   53.7     24  2  B27262             nicotinic acetylch
    43      22   53.7     25  2  A48363             2-hydroxyglutaryl-
    44      22   53.7     26  4  I55277             hemoglobin alpha c
    45      22   53.7     30  2  S55462             mer5 protein homol

                              ALIGNMENTS

RESULT 1
T30514
hypothetical protein - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
C;Accession: T30514
R;Onai, K.; Katagiri, S.; Akiyama, M.; Nakashima, H.
Mol. Gen. Genet. 259, 264-271, 1998
A;Title: Mutation of the gene for the second-largest subunit of RNA polymerase I prolong
A;Reference number: Z20838; MUID:98420098; PMID:9749669
A;Accession: T30514
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-30 <ONA>
A;Cross-references: EMBL:AB006052; NID:g3668169; PIDN:BAA33444.1; PID:g3668170

  Query Match             73.2%;  Score 30;  DB 2;  Length 30;
  Best Local Similarity   83.3%;  Pred. No. 36;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPA 6
              |||:||
Db         11 PDTQPA 16


RESULT 2
S44201
HLA-DRB1 exon2 protein - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Jan-1995
C;Accession: S44201
R;Loeffler, D.; Kaltenbacher, U.; Woelpl, A.; Eiermann, T.H.
submitted to the EMBL Data Library, April 1994
A;Description: Nucleotide sequence of a novel HLA-DRB1 allele, DRB1*08??.
A;Reference number: S44201
A;Accession: S44201
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-25 <LOE>
A;Cross-references: EMBL:Z32685

  Query Match             68.3%;  Score 28;  DB 2;  Length 25;
  Best Local Similarity   71.4%;  Pred. No. 69;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              | | |||
Db         17 PGTAPAP 23


RESULT 3
B60745
```

```
major outer membrane protein - Fusobacterium nucleatum (ATCC 25586) (fragment)
C;Species: Fusobacterium nucleatum
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: B60745
R;Bakken, V.; Aaro, S.; Jensen, H.B.
J. Gen. Microbiol. 135, 3253-3262, 1989
A;Title: Purification and partial characterization of a major outer-membrane protein of
A;Reference number: A60745; MUID:90257576; PMID:2636259
A;Accession: B60745
A;Molecule type: protein
A;Residues: 1-26 <BAK>
C;Keywords: membrane protein

  Query Match             65.9%;  Score 27;  DB 2;  Length 26;
  Best Local Similarity   71.4%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |  ||||
Db          6 PAPRPAP 12


RESULT 4
A30114
long neurotoxin I - sea anemone (Stichodactyla helianthus)
C;Species: Stichodactyla helianthus, Stoichactis helianthus (Caribbean sea anemone)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A30114; S47619
R;Kem, W.R.; Parten, B.; Pennington, M.W.; Price, D.A.; Dunn, B.M.
Biochemistry 28, 3483-3489, 1989
A;Title: Isolation, characterization, and amino acid sequence of a polypeptide neurotoxi
A;Reference number: A30114; MUID:89302949; PMID:2568126
A;Accession: A30114
A;Molecule type: protein
A;Residues: 1-48 <KEM>
A;Cross-references: UNIPROT:P19651
R;Monks, S.A.; Gould, A.R.; Lumley, P.E.; Alewood, P.F.; Kem, W.R.; Goss, N.H.; Norton,
Biochim. Biophys. Acta 1207, 93-101, 1994
A;Title: Limited proteolysis study of structure-function relationships in Sh I, a polype
A;Reference number: S47619; MUID:94318675; PMID:7913833
A;Accession: S47619
A;Molecule type: protein
A;Residues: 1-48 <MON>
C;Superfamily: sea anemone toxin
C;Keywords: neurotoxin

  Query Match             65.9%;  Score 27;  DB 1;  Length 48;
  Best Local Similarity   71.4%;  Pred. No. 2e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              || | ||
Db         10 PDIRTAP 16


RESULT 5
S38295
serine proteinase - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 18-Feb-1994 #sequence_revision 17-Nov-1995 #text_change 09-Jul-2004
C;Accession: S38295
R;Heidtmann, H.H.; Travis, J.
Biol. Chem. Hoppe-Seyler 374, 871-875, 1993
A;Title: A novel chymotrypsin-like serine proteinase from human lung.
A;Reference number: S38295; MUID:94092341; PMID:8267879
A;Accession: S38295
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-26 <HEI>
A;Cross-references: UNIPROT:P34168
C;Superfamily: trypsin; trypsin homology

  Query Match             63.4%;  Score 26;  DB 2;  Length 26;
```

```
  Best Local Similarity   80.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRP 5
              ||:||
Db          9 PDSRP 13


RESULT 6
C95037
hypothetical protein SP0316 [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: C95037
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heic
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: C95037
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-35 <KUR>
A;Cross-references: UNIPROT:Q97SL6; GB:AE005672; PIDN:AAK74492.1; PID:g14971789; GSPDB:(
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0316

  Query Match             61.0%;  Score 25;  DB 2;  Length 35;
  Best Local Similarity   57.1%;  Pred. No. 3.4e+02;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |: :|||
Db          9 PNHQPAP 15


RESULT 7
I70082
glycoprotein Ib alpha variant B - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-Jul-2004
C;Accession: I70082
R;Lopez, J.A.; Ludwig, E.H.; McCarthy, B.J.
J. Biol. Chem. 267, 10055-10061, 1992
A;Title: Polymorphism of human glycoprotein Ib alpha results from a variable number of t
ations.
A;Reference number: I55355; MUID:92250564; PMID:1577776
A;Accession: I70082
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-42 <RES>
A;Cross-references: UNIPROT:Q16469; GB:S34439; NID:g249178; PIDN:AAB22153.1; PID:g24917
C;Superfamily: platelet glycoprotein Ib alpha chain; leucine-rich alpha-2-glycoprotein
C;Keywords: glycoprotein

  Query Match             61.0%;  Score 25;  DB 2;  Length 42;
  Best Local Similarity   57.1%;  Pred. No. 4.1e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              | : |||
Db          2 PTSEPAP 8


RESULT 8
S07073
arabinogalactan protein - Italian ryegrass (fragments)
C;Species: Lolium multiflorum (Italian ryegrass)
C;Date: 21-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
C;Accession: S07073
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         December 29, 2004, 13:27:35 ; Search time 68 Seconds
                                          (without alignments)
                                          37.031 Million cell updates/sec

Title:          US-09-606-910E-3
Perfect score:  41
Sequence:       1 PDTRPAP 7

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters:      431440

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_AA:*
                1:  /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
                2:  /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
                3:  /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
                4:  /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
                5:  /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
                6:  /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
                7:  /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
                8:  /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
                9:  /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
                10:  /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
                11:  /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
                12:  /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
                13:  /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
                14:  /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
                15:  /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
                16:  /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
                17:  /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
                18:  /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
                19:  /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
                20:  /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                     %
Result             Query
   No.    Score    Match Length DB  ID                          Description
 -----------------------------------------------------------------------------
     1       41    100.0      7 14  US-10-057-136-3              Sequence 3, Appli
     2       41    100.0      8  9  US-09-994-466-2             Sequence 2, Appli
     3       41    100.0      8 15  US-10-296-317-43             Sequence 43, Appl
     4       41    100.0     12 13  US-10-101-731-1              Sequence 1, Appli
     5       41    100.0     13 14  US-10-080-608A-77            Sequence 77, Appl
     6       41    100.0     13 15  US-10-371-067-7              Sequence 7, Appli
     7       41    100.0     15  9  US-09-822-698A-40           Sequence 40, Appl
     8       41    100.0     16  9  US-09-984-183-17            Sequence 17, Appl
     9       41    100.0     16  9  US-09-984-333-7             Sequence 7, Appli
    10       41    100.0     16 11  US-09-143-379-1              Sequence 1, Appli
    11       41    100.0     18  9  US-09-815-346-3             Sequence 3, Appli
    12       41    100.0     18 14  US-10-106-876-6              Sequence 6, Appli
    13       41    100.0     18 14  US-10-106-876-17             Sequence 17, Appl
    14       41    100.0     19  9  US-09-815-346-2             Sequence 2, Appli
    15       41    100.0     19  9  US-09-815-346-6             Sequence 6, Appli
    16       41    100.0     19 14  US-10-106-876-5              Sequence 5, Appli
    17       41    100.0     20  9  US-09-847-185-20            Sequence 20, Appl
    18       41    100.0     20  9  US-09-994-466-1             Sequence 1, Appli
    19       41    100.0     20  9  US-09-984-183-11            Sequence 11, Appl
    20       41    100.0     20  9  US-09-984-183-16            Sequence 16, Appl
    21       41    100.0     20  9  US-09-822-698A-7            Sequence 7, Appli
    22       41    100.0     20  9  US-09-822-698A-8            Sequence 8, Appli
    23       41    100.0     20  9  US-09-984-333-1             Sequence 1, Appli
    24       41    100.0     20  9  US-09-984-333-6             Sequence 6, Appli
    25       41    100.0     20  9  US-09-984-333-9             Sequence 9, Appli
    26       41    100.0     20 10  US-09-834-240-1              Sequence 1, Appli
    27       41    100.0     20 13  US-10-054-488-1              Sequence 1, Appli
    28       41    100.0     20 14  US-10-057-136-1              Sequence 1, Appli
    29       41    100.0     20 14  US-10-224-286-20             Sequence 20, Appl
    30       41    100.0     20 14  US-10-335-394-40             Sequence 40, Appl
    31       41    100.0     20 14  US-10-106-876-19             Sequence 19, Appl
    32       41    100.0     20 14  US-10-292-896-1              Sequence 1, Appli
    33       41    100.0     20 15  US-10-406-317-31             Sequence 31, Appl
    34       41    100.0     20 15  US-10-612-090-3              Sequence 3, Appli
    35       41    100.0     20 15  US-10-296-317-45             Sequence 45, Appl
    36       41    100.0     20 15  US-10-297-168-31             Sequence 31, Appl
    37       41    100.0     20 16  US-10-716-293-215            Sequence 215, App
    38       41    100.0     21 14  US-10-062-710-196            Sequence 196, App
    39       41    100.0     21 14  US-10-062-710-207            Sequence 207, App
    40       41    100.0     21 16  US-10-380-927-1              Sequence 1, Appli
    41       41    100.0     22 14  US-10-106-876-18             Sequence 18, Appl
    42       41    100.0     24 14  US-10-292-896-2              Sequence 2, Appli
    43       41    100.0     24 15  US-10-296-317-62             Sequence 62, Appl
    44       41    100.0     25 10  US-09-881-339-3              Sequence 3, Appli
    45       41    100.0     25 14  US-10-292-896-9              Sequence 9, Appli

                                ALIGNMENTS


RESULT 1
US-10-057-136-3
; Sequence 3, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
;  APPLICANT: SCHLOM, JEFFREY
;  APPLICANT:  KANTOR, JUDITH
;  APPLICANT:  KUFE, DONALD
;  APPLICANT:  PANICALI, DENNIS
;  APPLICANT:  GRITZ, LINDA
;  TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
;  TITLE OF INVENTION:  TUMOR-ASSOCIATED ANTIGEN
;  FILE REFERENCE: 700953/47113C
;  CURRENT APPLICATION NUMBER: US/10/057,136
;  CURRENT FILING DATE:  2002-01-25
;  PRIOR APPLICATION NUMBER: 09/366,670
;  PRIOR FILING DATE: 1999-08-03
;  PRIOR APPLICATION NUMBER: PCT/US98/03693
;  PRIOR FILING DATE: 1998-02-24
;  PRIOR APPLICATION NUMBER: 60/038,253
;  PRIOR FILING DATE: 1997-02-24
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-057-136-3

  Query Match             100.0%;  Score 41;  DB 14;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          1 PDTRPAP 7
```

```
RESULT 2
US-09-994-466-2
; Sequence 2, Application US/09994466
; Publication No. US20020132771A1
; GENERAL INFORMATION:
;  APPLICANT: Madiyalakan, R.
;  TITLE OF INVENTION: THERAPEUTIC BINDING AGENTS AGAINST MUC-1 ANTIGEN AND METHODS FOR
;  TITLE OF INVENTION:  THEIR USE
;  FILE REFERENCE: AREX-P03-002
;  CURRENT APPLICATION NUMBER: US/09/994,466
;  CURRENT FILING DATE:  2001-11-26
;  PRIOR APPLICATION NUMBER: 09/724094
;  PRIOR FILING DATE: 2000-11-28
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
;   LENGTH: 8
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-994-466-2

  Query Match             100.0%;  Score 41;  DB 9;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          2 PDTRPAP 8


RESULT 3
US-10-296-317-43
; Sequence 43, Application US/10296317
; Publication No. US20040057968A1
; GENERAL INFORMATION:
;  APPLICANT: CEl-Sci Corp
;  APPLICANT:  Zimmerman, Daniel S
;  APPLICANT:  Sarin, Prem S
;  TITLE OF INVENTION: T CELL BINDING LIGAND PEPTIDES, PEPTIDE
;  FILE REFERENCE: CS-112
;  CURRENT APPLICATION NUMBER: US/10/296,317
;  CURRENT FILING DATE:  2002-11-22
;  PRIOR APPLICATION NUMBER: US 60/206548
;  PRIOR FILING DATE: 2000-05-24
;  PRIOR APPLICATION NUMBER: PCT/US07/16793
;  PRIOR FILING DATE: 2001-05-24
;  NUMBER OF SEQ ID NOS: 96
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43
;   LENGTH: 8
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Muc1 Peptide M1a
US-10-296-317-43

  Query Match             100.0%;  Score 41;  DB 15;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          2 PDTRPAP 8


RESULT 4
US-10-101-731-1
; Sequence 1, Application US/10101731
; Publication No. US20020168366A1
; GENERAL INFORMATION:
;  APPLICANT: Stewart, Michael
;  APPLICANT:  Person, Roland
;  APPLICANT:  No. US20020168366A1jaim, Antoine
;  TITLE OF INVENTION: Compositions and Methods for Producing Vascular Occlusion
;  FILE REFERENCE: T57005US
;  CURRENT APPLICATION NUMBER: US/10/101,731
;  CURRENT FILING DATE:  2002-03-21
;  PRIOR APPLICATION NUMBER: US/09/438,944A
;  PRIOR FILING DATE: 1999-11-12
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 12
;   TYPE: PRT
;   ORGANISM: human MUC-1
US-10-101-731-1

  Query Match             100.0%;  Score 41;  DB 13;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 3.8;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          6 PDTRPAP 12


RESULT 5
US-10-080-608A-77
; Sequence 77, Application US/10080608A
; Publication No. US20030198956A1
; GENERAL INFORMATION:
;  APPLICANT: Makowski, Lee
;  APPLICANT:  Hyman, Paul
;  APPLICANT:  Williams, Mark
;  TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
;  FILE REFERENCE: 8471-010-999
;  CURRENT APPLICATION NUMBER: US/10/080,608A
;  CURRENT FILING DATE:  2002-02-21
;  NUMBER OF SEQ ID NOS: 180
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 77
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: Derived from Homo sapiens MUC1 epithelial mucin glycoprotein.
US-10-080-608A-77

  Query Match             100.0%;  Score 41;  DB 14;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 4.1;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          4 PDTRPAP 10


RESULT 6
US-10-371-067-7
; Sequence 7, Application US/10371067
; Publication No. US20040018587A1
; GENERAL INFORMATION:
;  APPLICANT: Makowski, Lee
;  APPLICANT:  Williams, Mark K
;  APPLICANT:  Goldberg, Edward B
;  TITLE OF INVENTION: Nanostructures containing antibody assembly subunits
;  FILE REFERENCE: NANF.P-003
;  CURRENT APPLICATION NUMBER: US/10/371,067
;  CURRENT FILING DATE:  2003-02-21
;  PRIOR APPLICATION NUMBER: 10/080,608
;  PRIOR FILING DATE: 2002-02-21
;  PRIOR APPLICATION NUMBER: 10/136,225
;  PRIOR FILING DATE: 2002-04-29
;  PRIOR APPLICATION NUMBER: 09/236,949
```

```
                      GenCore version 5.1.6
             Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         December 29, 2004, 13:17:24 ; Search time 24 Seconds
                                             (without alignments)
                                             19.343 Million cell updates/sec

Title:          US-09-606-910E-3
Perfect score:  41
Sequence:       1 PDTRPAP 7

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters:      260538

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                2:  /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                3:  /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                4:  /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                5:  /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                6:  /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                         Description
-------------------------------------------------------------------------------
    1      41   100.0      8  4  US-09-641-833-2            Sequence 2, Appli
    2      41   100.0      9  2  US-08-288-059-13           Sequence 13, Appl
    3      41   100.0      9  2  US-08-288-059-27           Sequence 27, Appl
    4      41   100.0      9  2  US-08-666-473-110          Sequence 110, App
    5      41   100.0      9  3  US-08-134-198E-42          Sequence 42, Appl
    6      41   100.0      9  3  US-09-043-731-9            Sequence 9, Appli
    7      41   100.0      9  3  US-09-043-731-10           Sequence 10, Appl
    8      41   100.0      9  4  US-09-593-870A-8           Sequence 8, Appli
    9      41   100.0      9  4  US-09-593-870A-29          Sequence 29, Appl
   10      41   100.0      9  4  US-09-593-870A-69          Sequence 69, Appl
   11      41   100.0      9  4  US-09-497-232-2            Sequence 2, Appli
   12      41   100.0      9  4  US-09-497-232-19           Sequence 19, Appl
   13      41   100.0      9  4  US-09-497-232-20           Sequence 20, Appl
   14      41   100.0     10  3  US-09-339-944-9            Sequence 9, Appli
   15      41   100.0     10  4  US-09-497-232-28           Sequence 28, Appl
   16      41   100.0     10  4  US-09-497-232-29           Sequence 29, Appl
   17      41   100.0     10  4  US-09-497-232-30           Sequence 30, Appl
   18      41   100.0     10  4  US-09-497-232-31           Sequence 31, Appl
   19      41   100.0     10  4  US-09-651-265-9            Sequence 9, Appli
   20      41   100.0     13  4  US-09-914-259-77           Sequence 77, Appl
   21      41   100.0     14  2  US-08-463-230A-12          Sequence 12, Appl
   22      41   100.0     15  1  US-08-122-546-18           Sequence 18, Appl
   23      41   100.0     15  2  US-08-764-938-18           Sequence 18, Appl
   24      41   100.0     15  3  US-09-131-052-18           Sequence 18, Appl
   25      41   100.0     15  3  US-09-131-053A-18          Sequence 18, Appl
   26      41   100.0     16  3  US-09-339-944-7            Sequence 7, Appli
   27      41   100.0     16  3  US-08-737-896-1            Sequence 1, Appli
   28      41   100.0     16  3  US-09-043-731-18           Sequence 18, Appl
   29      41   100.0     16  4  US-09-497-232-15           Sequence 15, Appl
   30      41   100.0     16  4  US-09-651-265-12           Sequence 12, Appl
   31      41   100.0     16  5  PCT-US96-09951-1           Sequence 1, Appli
   32      41   100.0     17  3  US-08-766-521-2            Sequence 2, Appli
   33      41   100.0     17  3  US-08-755-413-2            Sequence 2, Appli
   34      41   100.0     19  1  US-08-099-354-3            Sequence 3, Appli
   35      41   100.0     19  1  US-08-099-354-5            Sequence 5, Appli
   36      41   100.0     19  2  US-08-288-059-9            Sequence 9, Appli
   37      41   100.0     19  2  US-08-288-059-11           Sequence 11, Appl
   38      41   100.0     19  4  US-09-593-870A-22          Sequence 22, Appl
   39      41   100.0     20  1  US-08-328-536-1            Sequence 1, Appli
   40      41   100.0     20  2  US-08-288-059-1            Sequence 1, Appli
   41      41   100.0     20  2  US-08-288-059-12           Sequence 12, Appl
   42      41   100.0     20  2  US-08-288-059-32           Sequence 32, Appl
   43      41   100.0     20  2  US-08-902-516-20           Sequence 20, Appl
   44      41   100.0     20  2  US-08-833-807-1            Sequence 1, Appli
   45      41   100.0     20  2  US-08-833-807-8            Sequence 8, Appli

                              ALIGNMENTS


RESULT 1
US-09-641-833-2
; Sequence 2, Application US/09641833
; Patent No. 6716966
; GENERAL INFORMATION:
;  APPLICANT: Madiyalakan, Ragupathy
;  TITLE OF INVENTION: Therapeutic Binding Agents Against MUC-1 Antigen and
;  TITLE OF INVENTION:  Methods
;  TITLE OF INVENTION:  for Their Use
;  FILE REFERENCE: 107823.127
;  CURRENT APPLICATION NUMBER: US/09/641,833
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 8
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: MUC-1 antigen
US-09-641-833-2

  Query Match             100.0%;  Score 41;  DB 4;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          2 PDTRPAP 8


RESULT 2
US-08-288-059-13
; Sequence 13, Application US/08288059
; Patent No. 5827666
;  GENERAL INFORMATION:
;    APPLICANT:  FINN, OLIVERA J.
;    APPLICANT:  FONTENOT, J. D.
;    APPLICANT:  MONTELARO, RONALD C.
;    TITLE OF INVENTION:  SYNTHETIC MULTIPLE TANDEM REPEAT MUCIN
;    TITLE OF INVENTION:  AND MUCIN-LIKE PEPTIDES, AND USES THEREOF
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN DARBY & CUSHMAN, L.L.P.
;      STREET:  1100 NEW YORK AVENUE, N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
```

```
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/288,059
;     FILING DATE:  08-AUG-1994
;     CLASSIFICATION:  424
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  CHAPIN, MARLANA K.
;     REGISTRATION NUMBER:  35,843
;     REFERENCE/DOCKET NUMBER:  61137/205204
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  202-861-3711
;     TELEFAX:  202-822-0944
;     TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  13:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  9 amino acids
;     TYPE:  amino acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  peptide
US-08-288-059-13

  Query Match             100.0%;  Score 41;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          1 PDTRPAP 7


RESULT 3
US-08-288-059-27
; Sequence 27, Application US/08288059
; Patent No. 5827666
;  GENERAL INFORMATION:
;    APPLICANT:  FINN, OLIVERA J.
;    APPLICANT:  FONTENOT, J. D.
;    APPLICANT:  MONTELARO, RONALD C.
;    TITLE OF INVENTION:  SYNTHETIC MULTIPLE TANDEM REPEAT MUCIN
;    TITLE OF INVENTION:  AND MUCIN-LIKE PEPTIDES, AND USES THEREOF
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN DARBY & CUSHMAN, L.L.P.
;      STREET:  1100 NEW YORK AVENUE, N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/288,059
;      FILING DATE:  08-AUG-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CHAPIN, MARLANA K.
;      REGISTRATION NUMBER:  35,843
;      REFERENCE/DOCKET NUMBER:  61137/205204
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-861-3711
;      TELEFAX:  202-822-0944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-288-059-27

  Query Match             100.0%;  Score 41;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          2 PDTRPAP 8


RESULT 4
US-08-666-473-110
; Sequence 110, Application US/08666473
; Patent No. 5843713
;  GENERAL INFORMATION:
;    APPLICANT:  YOSHIDA, Aruto
;    APPLICANT:  TAKEUCHI, Makoto
;    TITLE OF INVENTION:  PEPTIDE SEQUENCE THAT FORMS MUCIN SUGAR
;    TITLE OF INVENTION:  CHAIN AND TECHNIQUE FOR MODIFYING PROTEIN TO BE LINKED
;    TITLE OF INVENTION:  WITH MUCIN SUGAR CHAIN
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/666,473
;      FILING DATE:  19-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/JP95/02238
;      FILING DATE:  01-NOV-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-22101
;      FILING DATE:  09-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-269111
;      FILING DATE:  01-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16887/837
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  110:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-666-473-110

  Query Match             100.0%;  Score 41;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          1 PDTRPAP 7
```

```
                         GenCore version 5.1.6
                 Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          December 29, 2004, 13:13:04 ; Search time 74 Seconds
                                            (without alignments)
                                            33.934 Million cell updates/sec

Title:           US-09-606-910E-3
Perfect score:   41
Sequence:        1 PDTRPAP 7

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters:      900495

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        A_Geneseq_23Sep04:*
                 1:  geneseqp1980s:*
                 2:  geneseqp1990s:*
                 3:  geneseqp2000s:*
                 4:  geneseqp2001s:*
                 5:  geneseqp2002s:*
                 6:  geneseqp2003as:*
                 7:  geneseqp2003bs:*
                 8:  geneseqp2004s:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                               SUMMARIES
               %
Result       Query
  No.  Score Match Length DB  ID                 Description
-----------------------------------------------------------------
   1     41  100.0     7  2  AAY25108           Aay25108 Human MUC
   2     41  100.0     8  2  AAY25109           Aay25109 Mouse IgG
   3     41  100.0     8  4  AAB31286           Aab31286 Antigenic
   4     41  100.0     8  5  AAU82061           Aau82061 Antigenic
   5     41  100.0     8  5  AAE20933           Aae20933 Human tum
   6     41  100.0     8  8  ADL15165           Adl15165 Human tum
   7     41  100.0     9  2  AAR68021           Aar68021 Mucin pep
   8     41  100.0     9  2  AAR68007           Aar68007 Mucin pep
   9     41  100.0     9  2  AAW72723           Aaw72723 Mucin pep
  10     41  100.0     9  2  AAW72709           Aaw72709 Mucin pep
  11     41  100.0     9  2  AAY05217           Aay05217 MUC VNTR
  12     41  100.0     9  2  AAW67597           Aaw67597 T-cell ac
  13     41  100.0     9  2  AAW67584           Aaw67584 T-cell ac
  14     41  100.0     9  2  AAW67596           Aaw67596 T-cell ac
  15     41  100.0     9  4  AAB72499           Aab72499 Fusion pr
  16     41  100.0     9  4  AAE09568           Aae09568 Human muc
  17     41  100.0     9  4  AAE09520           Aae09520 Human muc
  18     41  100.0     9  4  AAE09519           Aae09519 Human muc
  19     41  100.0    10  2  AAW67608           Aaw67608 Human MUC
  20     41  100.0    10  2  AAW67609           Aaw67609 Human MUC
  21     41  100.0    10  2  AAW67606           Aaw67606 Human MUC
  22     41  100.0    10  2  AAW67607           Aaw67607 Human MUC
  23     41  100.0    10  3  AAY80107           Aay80107 Anti MUC-
  24     41  100.0    10  5  AAE20934           Aae20934 Human tum
  25     41  100.0    11  4  AAB49998           Aab49998 Tumour an
  26     41  100.0    11  5  AAE20935           Aae20935 Human tum
  27     41  100.0    12  5  AAE20936           Aae20936 Human tum
  28     41  100.0    13  2  AAY03870           Aay03870 SM3 antib
  29     41  100.0    13  7  ADC35694           Adc35694 Peptide e
  30     41  100.0    13  8  ADO15610           Ado15610 Human Bre
  31     41  100.0    14  2  AAR95561           Aar95561 Mucin cor
  32     41  100.0    14  2  AAY40237           Aay40237 Amino aci
  33     41  100.0    15  4  AAE12721           Aae12721 Human PH1
  34     41  100.0    15  5  AAE20939           Aae20939 MUC1 prot
  35     41  100.0    15  5  AAE20938           Aae20938 MUC1 prot
  36     41  100.0    15  5  AAE20955           Aae20955 MUC1 prot
  37     41  100.0    15  5  AAE20941           Aae20941 MUC1 prot
  38     41  100.0    15  5  AAE20942           Aae20942 MUC1 prot
  39     41  100.0    15  5  AAE20956           Aae20956 MUC1 prot
  40     41  100.0    15  5  AAE20943           Aae20943 MUC1 prot
  41     41  100.0    15  5  AAE20940           Aae20940 MUC1 prot
  42     41  100.0    15  5  AAE20957           Aae20957 MUC1 prot
  43     41  100.0    15  6  ABU58813           Abu58813 Mucin 1 (
  44     41  100.0    16  2  AAR84861           Aar84861 MUC-1 epi
  45     41  100.0    16  2  AAW21978           Aaw21978 Human MUC

                               ALIGNMENTS

RESULT 1
AAY25108
ID   AAY25108 standard; peptide; 7 AA.
XX
AC   AAY25108;
XX
DT   25-AUG-1999  (first entry)
XX
DE   Human MUC1 immunodominant epitope.
XX
KW   Immunodominant; epitope; MUC1; epithelial mucin; tumour; vaccine;
KW   mammary carcinoma; colorectal carcinoma; pancreas carcinoma.
XX
OS   Homo sapiens.
XX
PN   DE19758400-A1.
XX
PD   01-JUL-1999.
XX
PF   30-DEC-1997;   97DE-01058400.
XX
PR   30-DEC-1997;   97DE-01058400.
XX
PA   (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.
PA   (HANI/) HANISCH F.
XX
PI   Karsten U,  Hanisch F,  Paulsen H;
XX
DR   WPI; 1999-372374/32.
XX
PT   New tumour vaccine - for treating tumour cells of mammary tumours,
PT   colorectal and pancreas carcinomas.
XX
PS   Example 1; Page 2; 6pp; German.
XX
CC   This invention describes a novel tumour vaccine which comprises a
CC   synthetic peptide derived from human epithelial mucin MUC1. This active
CC   tumour vaccine can be used against tumour cells from mammary, colorectal
CC   or pancreas carcinomas. This sequence represents an immunodominant
CC   epitope from human MUC1
XX
SQ   Sequence 7 AA;

  Query Match             100.0%;  Score 41;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
```

```
              |||||||
Db          1 PDTRPAP 7


RESULT 2
AAY25109
ID   AAY25109 standard; peptide; 8 AA.
XX
AC   AAY25109;
XX
DT   25-AUG-1999  (first entry)
XX
DE   Mouse IgG1 epitope A76-A/C7.
XX
KW   Immunodominant; epitope; MUC1; epithelial mucin; tumour; vaccine;
KW   mammary carcinoma; colorectal carcinoma; pancreas carcinoma; murine;
KW   IgG1; A76-A/C7.
XX
OS   Mus sp.
XX
PN   DE19758400-A1.
XX
PD   01-JUL-1999.
XX
PF   30-DEC-1997;   97DE-01058400.
XX
PR   30-DEC-1997;   97DE-01058400.
XX
PA   (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.
PA   (HANI/) HANISCH F.
XX
PI   Karsten U,  Hanisch F,  Paulsen H;
XX
DR   WPI; 1999-372374/32.
XX
PT   New tumour vaccine - for treating tumour cells of mammary tumours,
PT   colorectal and pancreas carcinomas.
XX
PS   Example 1; Page 2; 6pp; German.
XX
CC   This invention describes a novel tumour vaccine which comprises a
CC   synthetic peptide derived from an immunodominant epitope from human
CC   epithelial mucin MUC1. This active tumour vaccine can be used against
CC   tumour cells from mammary, colorectal or pancreas carcinomas. This
CC   sequence represents a murine IgG1 derived epitope, A76-A/C7
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 41;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          2 PDTRPAP 8


RESULT 3
AAB31286
ID   AAB31286 standard; peptide; 8 AA.
XX
AC   AAB31286;
XX
DT   20-APR-2001  (first entry)
XX
DE   Antigenic peptide derived from a human MUC1 protein.
XX
KW   MUC1; epithelial cell mucin; tumour; MUC1 ligand; cancer.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO200077031-A2.
XX
PD   21-DEC-2000.
XX
PF   15-JUN-2000; 2000WO-CA000711.
XX
PR   15-JUN-1999;   99US-0139263P.
XX
PA   (UYHE-) UNIV HEALTH NETWORK.
XX
PI   Gariepy J,  Yang S;
XX
DR   WPI; 2001-091388/10.
XX
PT   Novel ligands that binds to MUC1, an epithelial cell mucin, useful for
PT   diagnosing, monitoring, treating and preventing cancer.
XX
PS   Example 2; Page 27; 60pp; English.
XX
CC   The present sequence represents a peptide derived from a human MUC1
CC   protein. MUC1 is an epithelial cell mucin, which is found on, and shed
CC   from, the surface of many tumours. The specification describes ligands
CC   that bind to MUC1. These ligands were isolated using a phage display
CC   technique using MUC1 tandem repeats as the target. The MUC1 tandem target
CC   repeats comprise repeats of the peptide AAB31258. The ligands are useful
CC   for detecting the presence of MUC1 in a sample, to treat or prevent
CC   cancer associated with MUC1 and to prepare a medicament or diagnostic
CC   agent to treat, prevent or detect cancer associated with MUC1. By
CC   assaying for the binding between the MUC1 ligand and MUC1 in a sample
CC   diagnosis or monitoring of cancer can be carried out. The ligands may
CC   also be used to prepare antibodies
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 41;  DB 4;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          2 PDTRPAP 8


RESULT 4
AAU82061
ID   AAU82061 standard; peptide; 8 AA.
XX
AC   AAU82061;
XX
DT   09-APR-2002  (first entry)
XX
DE   Antigenic peptide M1a associated with cancer (Muc1).
XX
KW   T-cell binding ligand; TCBL; peptide G'; human MHC class II beta chain;
KW   peptide J; human beta-2-microglobulin; HIV-1; TCBL peptide construct;
KW   immunological disorder; immune response; human immunodeficiency virus;
KW   herpes simplex virus infection; HSV; malaria; tuberculosis; cancer; CEA;
KW   acquired immunodeficiency syndrome; AIDS; allergy; autoimmune disease;
KW   autoimmune myocarditis; cytostatic; antiinflammatory.
XX
OS   Synthetic.
XX
PN   WO200189286-A2.
XX
PD   29-NOV-2001.
XX
PF   24-MAY-2001; 2001WO-US016793.
XX
PR   24-MAY-2000; 2000US-0206548P.
XX
PA   (CELS-) CEL-SCI CORP.
XX
PI   Zimmerman DS,  Sarin PS;
XX
```

```
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.72 kb. Life Technologies catalog #:
                     11548-013"
ORIGIN

Alignment Scores:
Pred. No.:              2.25e+03       Length:        330
Score:                  41.00          Matches:       7
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     1              Gaps:          0

US-09-606-910E-3 (1-7) x AI925867 (1-330)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db        277 CCGGACACCAGGCCGGCCCCG 257

RESULT 4
AZ780422
LOCUS       AZ780422                 349 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0017I15R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0017I15 R, genomic survey sequence.
ACCESSION   AZ780422
VERSION     AZ780422.1  GI:12912067
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 349)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0017  row: I  column: 15
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 349.
FEATURES             Location/Qualifiers
     source          1. .349
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0017I15"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

Alignment Scores:
Pred. No.:              2.37e+03       Length:        349
Score:                  41.00          Matches:       7
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     8              Gaps:          0

US-09-606-910E-3 (1-7) x AZ780422 (1-349)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db        166 CCAGACACCAGGCCTGCTCCC 186

RESULT 5
BQ625927/c
LOCUS       BQ625927                 362 bp    mRNA    linear   EST 20-DEC-2002
DEFINITION  ph86g06.y1 Ostertagia ostertagi L3 SL1 TOPO v2 Ostertagia ostertagi
            cDNA 5', mRNA sequence.
ACCESSION   BQ625927
VERSION     BQ625927.1  GI:21653105
KEYWORDS    EST.
SOURCE      Ostertagia ostertagi
  ORGANISM  Ostertagia ostertagi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
            Trichostrongyloidea; Haemonchidae; Ostertagiinae; Ostertagia.
REFERENCE   1  (bases 1 to 362)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
            Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
            Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
            Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
            Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Claire Murphy and Dr. James McCarter
            at Washington University, St. Louis. Nematodes were provided by Dr.
            Louis Gasbarree of the USDA, Beltsville, MD
            (lgasbarr@anri.barc.usda.gov).
            Seq primer: SL1 primer.
FEATURES             Location/Qualifiers
     source          1. .362
                     /organism="Ostertagia ostertagi"
                     /mol_type="mRNA"
                     /db_xref="taxon:6317"
                     /dev_stage="third stage exsheathed larvae"
                     /lab_host="DH10B"
                     /clone_lib="Ostertagia ostertagi L3 SL1 TOPO v2"
                     /note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI;
                     Site_2: EcoRI; ***WARNING: Subsequent examination of these
                     samples has revealed the presence of an additional
                     Trichostrongyloidea cattle nematode, Cooperia oncophora.
                     Sequences in this library may derive from either
                     Ostertagia or Cooperia.*** The library was constructed by
                     Claire Murphy and Dr. James McCarter at Washington
```

```
                    University, St. Louis. Oligo(dT)-SL1 PCR based library.
                    Ostertagia ostertagi L3 cDNA PCR products of size >400
                    nucleotides containing SL1 on the 5' end and oligo(dT) on
                    the 3' end were non-directionally cloned into
                    pCRII-TOPO(Invitrogen) following the Topo TA cloning
                    protocol. Nematodes were provided by Dr. Louis Gasbarree
                    of the USDA, Beltsville, MD (lgasbarr@anri.barc.usda.gov).
                    Third stage exsheathed larvae were collected from 14 day
                    fecal-sphagnum moss cultures of Ostertagia eggs. The
                    larvae were recovered by overnight passage on a Baermann
                    apparatus, and then cleaned by passage through a 20 micron
                    nylon mesh. The larvae were then subjected to a treatment
                    with 1.25% chlorox to induce excystation. The larvae were
                    washed with 5 changes of PBS and then pelleted and snap
                    frozen in liquid nitrogen."
ORIGIN

Alignment Scores:
Pred. No.:              2.45e+03      Length:        362
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     5             Gaps:          0

US-09-606-910E-3 (1-7) x BQ625927 (1-362)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         98 CCGGACACGAGGCCGGCACCG 78

RESULT 6
BQ625736/c
LOCUS       BQ625736                 372 bp    mRNA    linear   EST 20-DEC-2002
DEFINITION  ph92a07.y1 Ostertagia ostertagi L3 SL1 TOPO v2 Ostertagia ostertagi
            cDNA 5', mRNA sequence.
ACCESSION   BQ625736
VERSION     BQ625736.1  GI:21652914
KEYWORDS    EST.
SOURCE      Ostertagia ostertagi
  ORGANISM  Ostertagia ostertagi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
            Trichostrongyloidea; Haemonchidae; Ostertagiinae; Ostertagia.
REFERENCE   1  (bases 1 to 372)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
            Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
            Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
            Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
            Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Claire Murphy and Dr. James McCarter
            at Washington University, St. Louis. Nematodes were provided by Dr.
            Louis Gasbarree of the USDA, Beltsville, MD
            (lgasbarr@anri.barc.usda.gov).
            Putative full length read
            The vector to vector length is 404
            Seq primer: SL1 primer.
FEATURES             Location/Qualifiers
     source          1..372
                     /organism="Ostertagia ostertagi"
                     /mol_type="mRNA"
                     /db_xref="taxon:6317"
                     /dev_stage="third stage exsheathed larvae"
```

```
                    /lab_host="DH10B"
                    /clone_lib="Ostertagia ostertagi L3 SL1 TOPO v2"
                    /note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI;
                    Site_2: EcoRI; ***WARNING: Subsequent examination of these
                    samples has revealed the presence of an additional
                    Trichostrongyloidea cattle nematode, Cooperia oncophora.
                    Sequences in this library may derive from either
                    Ostertagia or Cooperia.*** The library was constructed by
                    Claire Murphy and Dr. James McCarter at Washington
                    University, St. Louis. Oligo(dT)-SL1 PCR based library.
                    Ostertagia ostertagi L3 cDNA PCR products of size >400
                    nucleotides containing SL1 on the 5' end and oligo(dT) on
                    the 3' end were non-directionally cloned into
                    pCRII-TOPO(Invitrogen) following the Topo TA cloning
                    protocol. Nematodes were provided by Dr. Louis Gasbarree
                    of the USDA, Beltsville, MD (lgasbarr@anri.barc.usda.gov).
                    Third stage exsheathed larvae were collected from 14 day
                    fecal-sphagnum moss cultures of Ostertagia eggs. The
                    larvae were recovered by overnight passage on a Baermann
                    apparatus, and then cleaned by passage through a 20 micron
                    nylon mesh. The larvae were then subjected to a treatment
                    with 1.25% chlorox to induce excystation. The larvae were
                    washed with 5 changes of PBS and then pelleted and snap
                    frozen in liquid nitrogen."
ORIGIN

Alignment Scores:
Pred. No.:              2.51e+03      Length:        372
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     5             Gaps:          0

US-09-606-910E-3 (1-7) x BQ625736 (1-372)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         54 CCGGACACGAGGCCGGCACCA 34

RESULT 7
BQ100139/c
LOCUS       BQ100139                 376 bp    mRNA    linear   EST 10-APR-2002
DEFINITION  ph73a04.y1 Ostertagia ostertagi L4 SL1 TOPO v1 Ostertagia ostertagi
            cDNA, mRNA sequence.
ACCESSION   BQ100139
VERSION     BQ100139.1  GI:20133123
KEYWORDS    EST.
SOURCE      Ostertagia ostertagi
  ORGANISM  Ostertagia ostertagi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
            Trichostrongyloidea; Haemonchidae; Ostertagiinae; Ostertagia.
REFERENCE   1  (bases 1 to 376)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
            Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
            Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
            Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
            Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Claire Murphy and Dr. James McCarter
            at Washington University, St. Louis. Nematodes were provided by Dr.
            Louis Gasbarree of the USDA, Beltsville, MD
            (lgasbarr@anri.barc.usda.gov). DNA Sequencing by: Washington
```

```
            University Genome Sequencing Center

            Putative full length read
            The vector to vector length is 377
            Seq primer: SL1 primer.
FEATURES             Location/Qualifiers
     source          1..376
                     /organism="Ostertagia ostertagi"
                     /mol_type="mRNA"
                     /db_xref="taxon:6317"
                     /dev_stage="L4"
                     /lab_host="DH10B"
                     /clone_lib="Ostertagia ostertagi L4 SL1 TOPO v1"
                     /note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI;
                     Site_2: EcoRI; The library was constructed by Claire
                     Murphy and Dr. James McCarter at Washington University,
                     St. Louis. Oligo(dT)-SL1 PCR based library. Ostertagia
                     ostertagi L4 cDNA PCR products of size >400 nucleotides
                     containing SL1 on the 5' end and oligo(dT) on the 3' end
                     were non-directionally cloned into pCRII-TOPO(Invitrogen)
                     following the Topo TA cloning protocol. Nematodes were
                     provided by Dr. Louis Gasbarree of the USDA, Beltsville,
                     MD (lgasbarr@anri.barc.usda.gov). Late fourth stage larvae
                     were recovered 10 days after infection of young calves
                     with Ostertagia L3. The calves were killed and the abomasa
                     (gastric glands) were recovered and washed extensively and
                     then placed in aBaermann apparatus containing warm PBS for
                     approximately 4 hours. The L4 that migrated from the
                     tissues were washed extensively, pelleted and snap frozen
                     in liquid nitrogen."
ORIGIN

Alignment Scores:
Pred. No.:              2.54e+03      Length:        376
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     5             Gaps:          0

US-09-606-910E-3 (1-7) x BQ100139 (1-376)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db        108 CCGGACACGAGGCCGGCACCG 88

RESULT 8
BU493463/c
LOCUS       BU493463                 378 bp    mRNA    linear   EST 11-SEP-2002
DEFINITION  Nb_ADGR_D3A05_T7 Nippostrongylus brasiliensis adult (full length
            enriched) Nippostrongylus brasiliensis cDNA clone Nb_ADGR_D3A05 5'
            similar to Y67D8B.4 CE31662 status:Confirmed TR:Q95XT9, mRNA
            sequence.
ACCESSION   BU493463
VERSION     BU493463.1  GI:22789718
KEYWORDS    EST.
SOURCE      Nippostrongylus brasiliensis
  ORGANISM  Nippostrongylus brasiliensis
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
            Trichostrongyloidea; Heligmonellidae; Nippostrongylinae;
            Nippostrongylus.
REFERENCE   1  (bases 1 to 378)
  AUTHORS   Harcus,Y., Fernandez,C. and Maizels,R.M.
  TITLE     A survey of genes expressed by Nippostrongylus brasiliensis
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Maizels RM
            Institute of Cell, Animal and Population Biology
            University of Edinburgh
            Ashworth Labs, West Mains Road, Edinburgh EH93JT, Scotland
            Tel: +44 131 650 5511
            Fax: +44 131 650 5450
            Email: r.maizels@ed.ac.uk
            The library was prepared by Yvonne Harcus for Rick Maizels
            (Edinburgh University, UK). Sequencing was also performed by
            Yvonne. The sequence contained a PolyA tail (trimmed)
            PCR PRimers
            FORWARD: M13F
            BACKWARD: M13R
            Plate: D3  row: A  column: 05
            Seq primer: T7
            High quality sequence stop: 378.
FEATURES             Location/Qualifiers
     source          1..378
                     /organism="Nippostrongylus brasiliensis"
                     /mol_type="mRNA"
                     /db_xref="taxon:27835"
                     /clone="Nb_ADGR_D3A05"
                     /sex="mixed"
                     /dev_stage="adult worm (day 6 post infection)"
                     /clone_lib="Nippostrongylus brasiliensis adult (full
                     length enriched)"
                     /note="Vector: pSPORT 1; Site_1: SalI; Site_2: NotI;
                     Nippostrongylus brasiliensis is a nematode parasite of
                     rodents (definitive host, rat) widely used in
                     immunological research. The library was constructed from
                     adult worms taken from the small intestine, 6 days after
                     subcutaneous inoculation of infective larvae. cDNA was
                     prepared by RT-PCR and directionally cloned in a plasmid
                     vector.  RT was primed with a tagged oligo-dT (NotI
                     primer-adapter). This tag was  then used to prime cDNA
                     amplification together with a primer directed  against an
                     oligonucleotide previously ligated to the originally
                     capped  mRNA population. cDNA was ligated to SalI
                     adapters, digested with NotI,  and cloned into pSPORT 1
                     cut with SalI and NotI."
ORIGIN

Alignment Scores:
Pred. No.:              2.55e+03      Length:        378
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     5             Gaps:          0

US-09-606-910E-3 (1-7) x BU493463 (1-378)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db        129 CCGGACACGAGACCAGCGCCA 109

RESULT 9
CL536602/c
LOCUS       CL536602                 381 bp    DNA     linear   GSS 20-MAY-2004
DEFINITION  Bma00046 B. mallei C-5 specific library of differential DNA
            fragments Burkholderia mallei genomic clone 1265, genomic survey
            sequence.
ACCESSION   CL536602
VERSION     CL536602.1  GI:47531183
KEYWORDS    GSS.
SOURCE      Burkholderia mallei
  ORGANISM  Burkholderia mallei
            Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
            Burkholderiaceae; Burkholderia.
REFERENCE   1  (bases 1 to 381)
  AUTHORS   Fushan,A., Monastyrskaya,G., Abaev,I., Filyukova,O., Kostina,M.,
            Pecherskih,E. and Sverdlov,E.
  TITLE     Genome-wide identification and mapping differences between the
            genomes of B. mallei and B. pseudomallei
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Fushan AA
            Laboratory of Structure and Functions of Human Genes
            Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of
            Sciences
            16/10 Miklukho-Maklaya St., 117997 GSP, Moscow V-437, Russia
            Tel: +7 095 330 6538
```

```
             Fax: +7 095 330 6538
             Email: gal@humgen.siobc.ras.ru
             B. mallei C-5 genomic sequence absent from B. pseudomallei C-141,
             misc_feature
             Seq primer: M13 Forward
             Class: plasmid insert
             High quality sequence stop: 381.
FEATURES             Location/Qualifiers
     source          1. .381
                     /organism="Burkholderia mallei"
                     /mol_type="genomic DNA"
                     /strain="C-5"
                     /db_xref="taxon:13373"
                     /clone="1265"
                     /clone_lib="B. mallei C-5 specific library of differential
                     DNA fragments"
                     /note="The library of B. mallei C-5 specific DNA fragments
                     absent from B. pseudomallei C-141 genome was obtained
                     using  a version of subtractive hybridization method
                     described in  [Schmidt KD, Schmidt-Rose T, Romling U,
                     Tummler B, Differential  genome analysis of bacteria by
                     genomic subtractive hybridization  and pulsed field gel
                     electrophoresis. Electrophoresis. 1998  Apr;19(4):509-14.]
                     with modifications to avoid driver  biotinilation. DNAs
                     from B. mallei C-5 and B. pseudomallei C-141  were used as
                     tester and driver, respectively."
ORIGIN

Alignment Scores:
Pred. No.:              2.57e+03       Length:        381
Score:                  41.00          Matches:       7
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     9              Gaps:          0

US-09-606-910E-3 (1-7) x CL536602 (1-381)

Qy          1 ProAspThrArgProAlaPro 7
                |||||||||||||||||||||
Db        238 CCAGATACACGACCGGCACCA 218

RESULT 10
BQ787472/c
LOCUS       BQ787472                 385 bp    mRNA    linear   EST 26-JUL-2002
DEFINITION  im10f10.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6034722
            5' similar to TR:Q9ULD0 Q9ULD0 KIAA1290 PROTEIN ;, mRNA sequence.
ACCESSION   BQ787472
VERSION     BQ787472.1  GI:21995944
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 385)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. J. Ferrer In vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
```

```
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 378.
FEATURES             Location/Qualifiers
     source          1. .385
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6034722"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue  following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact  Hiroshi Inoue, MD/PhD for further
                     information on this library  (Metabolism Division, Permutt
                     Laboratory, Washington University School  of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110).  Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN

Alignment Scores:
Pred. No.:              2.6e+03        Length:        385
Score:                  41.00          Matches:       7
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     5              Gaps:          0

US-09-606-910E-3 (1-7) x BQ787472 (1-385)

Qy          1 ProAspThrArgProAlaPro 7
                |||||||||||||||||||||
Db        118 CCTGACACCAGGCCAGCTCCG 98

RESULT 11
BQ626165/c
LOCUS       BQ626165                 396 bp    mRNA    linear   EST 20-DEC-2002
DEFINITION  ph94c06.y1 Ostertagia ostertagi L3 SL1 TOPO v2 Ostertagia ostertagi
            cDNA 5', mRNA sequence.
ACCESSION   BQ626165
VERSION     BQ626165.1  GI:21653343
KEYWORDS    EST.
SOURCE      Ostertagia ostertagi
  ORGANISM  Ostertagia ostertagi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
            Trichostrongyloidea; Haemonchidae; Ostertagiinae; Ostertagia.
REFERENCE   1  (bases 1 to 396)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
            Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
            Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
            Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
            Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Claire Murphy and Dr. James McCarter
            at Washington University, St. Louis. Nematodes were provided by Dr.
            Louis Gasbarree of the USDA, Beltsville, MD
            (lgasbarr@anri.barc.usda.gov).
```

```
            Seq primer: SL1 primer
            High quality sequence stop: 391.
FEATURES             Location/Qualifiers
     source          1. .396
                     /organism="Ostertagia ostertagi"
                     /mol_type="mRNA"
                     /db_xref="taxon:6317"
                     /dev_stage="third stage exsheathed larvae"
                     /lab_host="DH10B"
                     /clone_lib="Ostertagia ostertagi L3 SL1 TOPO v2"
                     /note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI;
                     Site_2: EcoRI; ***WARNING: Subsequent examination of these
                     samples has revealed the presence of an additional
                     Trichostrongyloidea cattle nematode, Cooperia oncophora.
                     Sequences in this library may derive from either
                     Ostertagia or Cooperia.*** The library was constructed by
                     Claire Murphy and Dr. James McCarter at Washington
                     University, St. Louis. Oligo(dT)-SL1 PCR based library.
                     Ostertagia ostertagi L3 cDNA PCR products of size >400
                     nucleotides containing SL1 on the 5' end and oligo(dT) on
                     the 3' end were non-directionally cloned into
                     pCRII-TOPO(Invitrogen) following the Topo TA cloning
                     protocol. Nematodes were provided by Dr. Louis Gasbarree
                     of the USDA, Beltsville, MD (lgasbarr@anri.barc.usda.gov).
                     Third stage exsheathed larvae were collected from 14 day
                     fecal-sphagnum moss cultures of Ostertagia eggs. The
                     larvae were recovered by overnight passage on a Baermann
                     apparatus, and then cleaned by passage through a 20 micron
                     nylon mesh. The larvae were then subjected to a treatment
                     with 1.25% chlorox to induce excystation. The larvae were
                     washed with 5 changes of PBS and then pelleted and snap
                     frozen in liquid nitrogen."
ORIGIN

Alignment Scores:
Pred. No.:              2.67e+03       Length:          396
Score:                  41.00          Matches:         7
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            100.00%        Indels:          0
DB:                     5              Gaps:            0

US-09-606-910E-3 (1-7) x BQ626165 (1-396)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db        100 CCGGACACGAGGCCGGCACCG 80

RESULT 12
BF882322
LOCUS       BF882322                 408 bp    mRNA    linear   EST 17-JAN-2001
DEFINITION  IL3-ET0113-051200-372-F01 ET0113 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF882322
VERSION     BF882322.1  GI:12272448
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 408)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL3&t2=IL3-ET0113-
            051200-372-F01&t3=2000-12-05&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 407.
FEATURES             Location/Qualifiers
     source          1. .408
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="ET0113"
                     /note="Organ: lung_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:              2.74e+03       Length:          408
Score:                  41.00          Matches:         7
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            100.00%        Indels:          0
DB:                     2              Gaps:            0

US-09-606-910E-3 (1-7) x BF882322 (1-408)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db        191 CCGGATACCCGACCAGCCCCA 211

RESULT 13
CB043230/c
LOCUS       CB043230                 411 bp    mRNA    linear   EST 16-JAN-2003
DEFINITION  Tc_L3_19H08_SKPL Teladorsagia circumcincta L3 library Teladorsagia
            circumcincta cDNA clone Tc_L3_19H08 5' similar to Q95XT9
            Hypothetical protein Y67D8B.4. Caenorhabditis elegans, mRNA
            sequence.
ACCESSION   CB043230
VERSION     CB043230.1  GI:27770385
KEYWORDS    EST.
SOURCE      Teladorsagia circumcincta
  ORGANISM  Teladorsagia circumcincta
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
            Trichostrongyloidea; Haemonchidae; Ostertagiinae; Teladorsagia.
REFERENCE   1  (bases 1 to 411)
  AUTHORS   Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Guiliano,D.,
            Hall,N., Quayle,M. and Barrell,B.
  TITLE     Edinburgh University/Sanger Centre Nematode EST Project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Blaxter ML
            Institute of Cell, Animal and Population Biology
            University of Edinburgh
            Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
            3JT, UK.
            Tel: +44 131 650 6760
            Fax: +44 131 670 5450
            Email: mark.blaxter@ed.ac.uk
            The library was prepared by Dr Diane Redmond and Dr David Knox,
            Moredun Research Institute, Midlothian, UK.
            PCR was performed by Ye Jieru, ICAPB, University of Edinburgh.
            Sequencing was performed by the Pathogen Sequencing Unit, Sanger
```

```
            Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).
            PCR PRimers
            FORWARD: T3
            BACKWARD: T7PL
            Plate: 19  row: H  column: 08
            Seq primer: SKPL
            High quality sequence stop: 411.
FEATURES             Location/Qualifiers
     source          1. .411
                     /organism="Teladorsagia circumcincta"
                     /mol_type="mRNA"
                     /db_xref="taxon:45464"
                     /clone="Tc_L3_19H08"
                     /sex="mixed"
                     /dev_stage="L3"
                     /clone_lib="Teladorsagia circumcincta L3 library"
                     /note="Vector: Lambda Uni-ZAP; Site_1: EcoRI; Site_2:
                     XhoI; Teladorsagia circumcincta is a parasitic nematode.
                     The library was constructed from mRNA from Teladorsagia
                     circumcincta L3 stage."
ORIGIN

Alignment Scores:
Pred. No.:              2.76e+03        Length:         411
Score:                  41.00           Matches:        7
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-606-910E-3 (1-7) x CB043230 (1-411)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db        142 CCGGACACGAGGCCTGCACCG 122

RESULT 14
BU070786/c
LOCUS       BU070786               453 bp    mRNA    linear   EST 27-AUG-2002
DEFINITION  im25c02.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6036075
            5' similar to TR:Q9ULD0 Q9ULD0 KIAA1290 PROTEIN ;, mRNA sequence.
ACCESSION   BU070786
VERSION     BU070786.1  GI:22511975
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 453)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: im25c02.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. J. Ferrer In vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 430.
FEATURES             Location/Qualifiers
```

```
     source          1. .453
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6036075"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue  following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact  Hiroshi Inoue, MD/PhD for further
                     information on this library  (Metabolism Division, Permutt
                     Laboratory, Washington University School  of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110).  Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN

Alignment Scores:
Pred. No.:              3.02e+03        Length:         453
Score:                  41.00           Matches:        7
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     5               Gaps:           0

US-09-606-910E-3 (1-7) x BU070786 (1-453)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         92 CCTGACACCAGGCCAGCTCCG 72

RESULT 15
CC821777/c
LOCUS       CC821777               468 bp    DNA     linear   GSS 01-JAN-2004
DEFINITION  MBSP7C3 Uncultured virus from nearshore marine surface sediment
            uncultured marine virus genomic, genomic survey sequence.
ACCESSION   CC821777
VERSION     CC821777.1  GI:40540166
KEYWORDS    GSS.
SOURCE      uncultured marine virus
  ORGANISM  uncultured marine virus
            Viruses; environmental samples.
REFERENCE   1  (bases 1 to 468)
  AUTHORS   Breitbart,M., Felts,B., Kelley,S., Mahaffy,J.M., Nulton,J.,
            Salamon,P. and Rohwer,F.
  TITLE     Diversity and Population Structure of a Nearshore Marine Sediment
            Viral Community
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rohwer F
            Biology Dept.
            San Diego State University
            5500 Campanile Dr, San Diego, CA 92102, USA
            Tel: 6195941336
            Fax: 619595676
            Email: forest@sunstroke.sdsu.edu
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .468
                     /organism="uncultured marine virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:186617"
                     /clone_lib="Uncultured virus from nearshore marine surface
                     sediment"
ORIGIN

Alignment Scores:
Pred. No.:              3.12e+03        Length:         468
Score:                  41.00           Matches:        7
Percent Similarity:     100.00%         Conservative:   0
```

```
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     9              Gaps:          0

US-09-606-910E-3 (1-7) x CC821777 (1-468)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db        145 CCAGATACCCGGCCGGCTCCC 125


Search completed: December 29, 2004, 14:55:20
Job time : 2052 secs
```

```
Alignment Scores:
Pred. No.:               28.3          Length:        60
Score:                   41.00         Matches:       7
Percent Similarity:      100.00%       Conservative:  0
Best Local Similarity:   100.00%       Mismatches:    0
Query Match:             100.00%       Indels:        0
DB:                      17            Gaps:          0

US-09-606-910E-3 (1-7) x US-10-716-293-214 (1-60)

Qy           1 ProAspThrArgProAlaPro 7
               |||||||||||||||||||||
Db           1 CCAGATACAAGACCCGCCCCT 21

RESULT 12
US-10-471-607-3
; Sequence 3, Application US/10471607
; Publication No. US20040115740A1
; GENERAL INFORMATION:
;  APPLICANT: The Victoria University of Manchester
;  APPLICANT:  Benson, Roderick
;  TITLE OF INVENTION: Intracellular analysis.
;  FILE REFERENCE: P088857PWO
;  CURRENT APPLICATION NUMBER: US/10/471,607
;  CURRENT FILING DATE:  2003-09-24
;  PRIOR APPLICATION NUMBER: GB 0108165.2
;  PRIOR FILING DATE: 3001-03-21
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 93
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: Artificial epitope construct
US-10-471-607-3

Alignment Scores:
Pred. No.:               40.8          Length:        93
Score:                   41.00         Matches:       7
Percent Similarity:      100.00%       Conservative:  0
Best Local Similarity:   100.00%       Mismatches:    0
Query Match:             100.00%       Indels:        0
DB:                      17            Gaps:          0

US-09-606-910E-3 (1-7) x US-10-471-607-3 (1-93)

Qy           1 ProAspThrArgProAlaPro 7
               |||||||||||||||||||||
Db          17 CCTGACACCAGACCTGCCCCT 37

RESULT 13
US-10-471-607-4/c
; Sequence 4, Application US/10471607
; Publication No. US20040115740A1
; GENERAL INFORMATION:
;  APPLICANT: The Victoria University of Manchester
;  APPLICANT:  Benson, Roderick
;  TITLE OF INVENTION: Intracellular analysis.
;  FILE REFERENCE: P088857PWO
;  CURRENT APPLICATION NUMBER: US/10/471,607
;  CURRENT FILING DATE:  2003-09-24
;  PRIOR APPLICATION NUMBER: GB 0108165.2
;  PRIOR FILING DATE: 3001-03-21
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 93
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: Artificial epitope construct
US-10-471-607-4

Alignment Scores:
Pred. No.:               40.8          Length:        93
Score:                   41.00         Matches:       7
Percent Similarity:      100.00%       Conservative:  0
Best Local Similarity:   100.00%       Mismatches:    0
Query Match:             100.00%       Indels:        0
DB:                      17            Gaps:          0

US-09-606-910E-3 (1-7) x US-10-471-607-4 (1-93)

Qy           1 ProAspThrArgProAlaPro 7
               |||||||||||||||||||||
Db          78 CCGGACACAAGGCCCGCCCCA 58

RESULT 14
US-10-259-678-644
; Sequence 644, Application US/10259678
; Publication No. US20030198974A1
; GENERAL INFORMATION:
;  APPLICANT: Cole, Stewart
;  APPLICANT:  Buchrieser-Brosch, Roland
;  APPLICANT:  Gordon, Stephen
;  APPLICANT:  Billault, Alain
;  TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
;  TITLE OF INVENTION:  THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
;  TITLE OF INVENTION:  LIBRARY APPLICATION TO THE DETECTION OF MYCOBACTERIA
;  FILE REFERENCE: 3495-0169
;  CURRENT APPLICATION NUMBER: US/10/259,678
;  CURRENT FILING DATE:  2002-09-30
;  PRIOR APPLICATION NUMBER: US/09/060,756
;  PRIOR FILING DATE: 1998-04-16
;  NUMBER OF SEQ ID NOS: 743
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 644
;   LENGTH: 146
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
US-10-259-678-644

Alignment Scores:
Pred. No.:               59.5          Length:        146
Score:                   41.00         Matches:       7
Percent Similarity:      100.00%       Conservative:  0
Best Local Similarity:   100.00%       Mismatches:    0
Query Match:             100.00%       Indels:        0
DB:                      15            Gaps:          0

US-09-606-910E-3 (1-7) x US-10-259-678-644 (1-146)

Qy           1 ProAspThrArgProAlaPro 7
               |||||||||||||||||||||
Db           7 CCTGATACTCGACCAGCCCCA 27

RESULT 15
US-10-471-607-5
; Sequence 5, Application US/10471607
; Publication No. US20040115740A1
; GENERAL INFORMATION:
;  APPLICANT: The Victoria University of Manchester
;  APPLICANT:  Benson, Roderick
;  TITLE OF INVENTION: Intracellular analysis.
;  FILE REFERENCE: P088857PWO
;  CURRENT APPLICATION NUMBER: US/10/471,607
;  CURRENT FILING DATE:  2003-09-24
;  PRIOR APPLICATION NUMBER: GB 0108165.2
;  PRIOR FILING DATE: 3001-03-21
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 156
;   TYPE: DNA
;   ORGANISM: Artificial
```

```
;   FEATURE:
;   OTHER INFORMATION: Artificial epitope construct
US-10-471-607-5

Alignment Scores:
Pred. No.:              62.8          Length:        156
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     17            Gaps:          0

US-09-606-910E-3 (1-7) x US-10-471-607-5 (1-156)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         17 CCTGACACCAGACCTGCCCCT 37


Search completed: December 29, 2004, 16:16:48
Job time : 1814 secs
```

```
Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db        517 CCGGACACCAGGCCGGCCCCG 537

RESULT 6
US-08-998-416-514/c
; Sequence 514, Application US/08998416
; Patent No. 6239264
;  GENERAL INFORMATION:
;    APPLICANT:  Philippsen, Peter
;    APPLICANT:  Pohlmann, Rainer
;    APPLICANT:  Steiner, Sabine
;    APPLICANT:  Mohr, Christine
;    APPLICANT:  Wendland, Jurgen
;    APPLICANT:  Knechtle, Philipp
;    APPLICANT:  Rebischung, Corinne
;    TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  1152
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6239264artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  No. 6239264th Carolina
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/998,416
;      FILING DATE:  24-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  CH 0016/97
;      FILING DATE:  31-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  PF/5-30306/A/CGC1976
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8587
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  514:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  854 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  PAG1362UP
US-08-998-416-514

Alignment Scores:
Pred. No.:              172            Length:        854
Score:                  41.00          Matches:       7
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     3              Gaps:          0

US-09-606-910E-3 (1-7) x US-08-998-416-514 (1-854)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db        450 CCGGATACCCGCCCAGCGCCA 430

RESULT 7
US-10-029-517-3
; Sequence 3, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
;  APPLICANT: Kenneth W. Dobie
;  APPLICANT:  Susan J. Myers
;  TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
;  FILE REFERENCE: RTS-0352
;  CURRENT APPLICATION NUMBER: US/10/029,517
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 3
;   LENGTH: 1721
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (58)...(1605)
US-10-029-517-3

Alignment Scores:
Pred. No.:              351            Length:        1721
Score:                  41.00          Matches:       7
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     4              Gaps:          0

US-09-606-910E-3 (1-7) x US-10-029-517-3 (1-1721)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db        601 CCGGACACCAGGCCGGCCCCG 621

RESULT 8
US-09-252-991A-10827
; Sequence 10827, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 10827
;   LENGTH: 1761
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10827

Alignment Scores:
Pred. No.:              359            Length:        1761
Score:                  41.00          Matches:       7
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     4              Gaps:          0

US-09-606-910E-3 (1-7) x US-09-252-991A-10827 (1-1761)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db       1634 CCCGATACCCGGCCTGCACCC 1654

RESULT 9
US-10-029-517-17
; Sequence 17, Application US/10029517
; Patent No. 6716627
```

```
; GENERAL INFORMATION:
;  APPLICANT: Kenneth W. Dobie
;  APPLICANT:  Susan J. Myers
;  TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
;  FILE REFERENCE: RTS-0352
;  CURRENT APPLICATION NUMBER: US/10/029,517
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 17
;   LENGTH: 1804
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (73)...(1500)
US-10-029-517-17

Alignment Scores:
Pred. No.:                368           Length:          1804
Score:                    41.00         Matches:         7
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       4             Gaps:            0

US-09-606-910E-3 (1-7) x US-10-029-517-17 (1-1804)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db        496 CCGGACACCAGGCCGGCCCCG 516

RESULT 10
US-09-252-991A-10853/c
; Sequence 10853, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 10853
;   LENGTH: 2256
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10853

Alignment Scores:
Pred. No.:                462           Length:          2256
Score:                    41.00         Matches:         7
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       4             Gaps:            0

US-09-606-910E-3 (1-7) x US-09-252-991A-10853 (1-2256)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db       1529 CCCGATACCCGGCCTGCACCC 1509

RESULT 11
US-09-620-312D-74/c
; Sequence 74, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
```

```
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 74
;   LENGTH: 3718
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (90)..(3122)
US-09-620-312D-74

Alignment Scores:
Pred. No.:                770           Length:          3718
Score:                    41.00         Matches:         7
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       4             Gaps:            0

US-09-606-910E-3 (1-7) x US-09-620-312D-74 (1-3718)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db       2919 CCTGACACCAGGCCAGCTCCG 2899

RESULT 12
US-10-029-517-19
; Sequence 19, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
;  APPLICANT: Kenneth W. Dobie
;  APPLICANT:  Susan J. Myers
;  TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
;  FILE REFERENCE: RTS-0352
;  CURRENT APPLICATION NUMBER: US/10/029,517
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 19
;   LENGTH: 8186
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 6899
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 7155
;   OTHER INFORMATION: unknown
```

```
;   NAME/KEY: unsure
;   LOCATION: 7184
;   OTHER INFORMATION: unknown
;   NAME/KEY: unsure
;   LOCATION: 7957
;   OTHER INFORMATION: unknown
;   NAME/KEY: intron
;   LOCATION: (2997)...(3498)
;   OTHER INFORMATION: intron 1
;   NAME/KEY: intron:exon junction
;   LOCATION: (3498)...(3499)
;   OTHER INFORMATION: intron 1:exon 2
;   NAME/KEY: exon
;   LOCATION: (3508)...(3599)
;   OTHER INFORMATION: exon 2d
;   NAME/KEY: exon:intron junction
;   LOCATION: (3982)...(3983)
;   OTHER INFORMATION: exon 2a:intron 2a
;   NAME/KEY: intron:exon junction
;   LOCATION: (4205)...(4206)
;   OTHER INFORMATION: intron 2c:exon 3c
;   NAME/KEY: intron:exon junction
;   LOCATION: (4259)...(4260)
;   OTHER INFORMATION: intron 2d:exon 3d
;   NAME/KEY: exon
;   LOCATION: (4260)...(4328)
;   OTHER INFORMATION: exon 3d
;   NAME/KEY: intron:exon junction
;   LOCATION: (4632)...(4633)
;   OTHER INFORMATION: intron 3:exon 4
;   NAME/KEY: exon
;   LOCATION: (4914)...(5035)
;   OTHER INFORMATION: exon 5
;   NAME/KEY: intron
;   LOCATION: (5266)...(6293)
;   OTHER INFORMATION: intron 6
US-10-029-517-19

Alignment Scores:
Pred. No.:              1.72e+03      Length:         8186
Score:                  41.00         Matches:        7
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     4             Gaps:           0

US-09-606-910E-3 (1-7) x US-10-029-517-19 (1-8186)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db       3864 CCGGACACCAGGCCGGCCCCG 3884

RESULT 13
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4403765
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: CDC 1551
;   OTHER INFORMATION: "n" bases at various positions throughout the sequence
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.:              7.78e+05      Length:         4403765
Score:                  41.00         Matches:        7
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     3             Gaps:           0

US-09-606-910E-3 (1-7) x US-09-103-840A-2 (1-4403765)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db     375904 CCTGATACTCGACCAGCCCCA 375884

RESULT 14
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.:              7.79e+05      Length:         4411529
Score:                  41.00         Matches:        7
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     3             Gaps:           0

US-09-606-910E-3 (1-7) x US-09-103-840A-1 (1-4411529)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db     375848 CCTGATACTCGACCAGCCCCA 375828

RESULT 15
US-09-252-991A-13024/c
; Sequence 13024, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 13024
```

```
;   LENGTH: 1149
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13024

Alignment Scores:
Pred. No.:              734           Length:        1149
Score:                  38.00         Matches:       6
Percent Similarity:     100.00%       Conservative:  1
Best Local Similarity:  85.71%        Mismatches:    0
Query Match:            92.68%        Indels:        0
DB:                     4             Gaps:          0

US-09-606-910E-3 (1-7) x US-09-252-991A-13024 (1-1149)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||:::|||
Db        160 CCAGATACCCGACCTTCTCCT 140


Search completed: December 29, 2004, 15:07:20
Job time : 720 secs
```

```
PD   31-OCT-1996.
XX
PF   28-APR-1995;   95DE-01015219.
XX
PR   28-APR-1995;   95DE-01015219.
XX
PA   (PECH/) PECHER G.
XX
PI   Pecher G;
XX
DR   WPI; 1996-486314/49.
DR   P-PSDB; AAW01455.
XX
PT   Assay for tumour-specific cytotoxic T cells - using stimulator and target
PT   B cells transfected with mucin gene sequence.
XX
PS   Claim 3; Fig 1; 5pp; German.
XX
CC   Stimulator and target cells are used in a novel assay to determine the
CC   specific cellular immune response of a patient to tumour-associated
CC   mucin. The stimulator and target cells are preferably Epstein-Barr virus-
CC   immortalised chimpanzee B cells that have been transfected with a human
CC   MUC1 gene sequence and treated with a glycosylation inhibitor. In
CC   particular, the MUC1 gene sequence includes 15-30 tandem repeats of the
CC   present 60 bp sequence
XX
SQ   Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              75.8          Length:        60
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     2             Gaps:          0

US-09-606-910E-3 (1-7) x AAT44506 (1-60)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db          1 CCGGACACCAGGCCGGCCCCG 21

RESULT 4
AAT38664
ID   AAT38664 standard; DNA; 60 BP.
XX
AC   AAT38664;
XX
DT   05-AUG-1997  (first entry)
XX
DE   Human MUC1 gene tandem repeat.
XX
KW   Tandem nucleotide repeat; mucin; epitope; vaccine; MUC1; tumour; breast;
KW   pancreas; ovary; colon; parotid gland; lung; therapy; plasmid pRep4MUC1;
KW   ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   repeat_unit     1. .60
FT                   /*tag=  a
FT                   /rpt_type= TANDEM
FT                   /transl_except= (pos:49. .51, aa:Val)
FT                   /note= "Preferably 15-30 copies are present in vectors
FT                   which are used for transfecting irradiated, EBV-
FT                   immortalised autologous human B-cells"
XX
PN   DE19516673-A1.
XX
PD   31-OCT-1996.
XX
PF   28-APR-1995;   95DE-01016673.
XX
```

```
PR   28-APR-1995;   95DE-01016673.
XX
PA   (PECH/) PECHER G.
XX
PI   Pecher G;
XX
DR   WPI; 1996-486530/49.
DR   P-PSDB; AAW04399.
XX
PT   Vaccines against human, pref. mucin expressing, tumours - comprising
PT   autologous B-cells transfected with mucin gene segments, treated with
PT   glycosylation inhibitor.
XX
PS   Claim 3; Fig 2; 5pp; German.
XX
CC   B-cells are isolated from the peripheral blood of a patient and are
CC   immortalised with Epstein-Barr virus (EBV) prior to transfection by
CC   electroporation with a vector containing segments of the human MUC1 gene.
CC   Preferably, there are 15-30 tandem repeats of the present 60 bp sequence
CC   in the vector. One specific vector was designated pRep4MUC1; it comprises
CC   the MUC1 sequences under the control of a RSV-LTR promoter, a hygromycin
CC   resistance gene under the control of a TK promoter and an EBNA-1 gene
CC   sequence. Cells are incubated for 24-48 hours in culture medium
CC   containing 5 mM of a glycosylation inhibitor. Finally the cells are
CC   irradiated with at least 10000 rad. The cells are useful as vaccines for
CC   therapy of mucin-expressing tumours, e.g. tumours of the breast,
CC   pancreas, ovary, colon, parotid gland and lung
XX
SQ   Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              75.8          Length:        60
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     2             Gaps:          0

US-09-606-910E-3 (1-7) x AAT38664 (1-60)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db          1 CCGGACACCAGGCCGGCCCCG 21

RESULT 5
AAT88166
ID   AAT88166 standard; DNA; 60 BP.
XX
AC   AAT88166;
XX
DT   28-APR-1998  (first entry)
XX
DE   Human MUC1 tandem nucleotide repeat.
XX
KW   MUC1; mucin; tandem repeat; anticancer vaccine; dendritic cell;
KW   tumour associated epitope; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .60
FT                   /*tag=  a
FT                   /transl_except= (pos:49. .51, aa: Val)
XX
PN   DE19617837-A1.
XX
PD   23-OCT-1997.
XX
PF   19-APR-1996;   96DE-01017837.
XX
PR   19-APR-1996;   96DE-01017837.
XX
PA   (PECH/) PECHER G.
```

```
XX
PI   Pecher G;
XX
DR   WPI; 1997-514604/48.
DR   P-PSDB; AAW32821.
XX
PT   Anticancer vaccine containing human dendritic cells - transfected with
PT   mucin gene containing tandem repeat sequences, and expressing tumour
PT   associated epitope when treated with glycosylation inhibitor.
XX
PS   Claim 4; Fig 2; 6pp; German.
XX
CC   This sequence represents a tandem repeat found in the MUC1 gene which
CC   encodes the glycoprotein mucin. This repeat region is used in a novel
CC   anticancer vaccine which also contains autologous human dendritic cells,
CC   capable of expressing tumour associated epitopes, preferably on the cell
CC   surface, when treated with a glycosylation inhibitor. The vaccine can be
CC   used to treat MUC1 expressing tumours, especially mammary, pancreatic,
CC   ovarian, colon and parotid tumours
XX
SQ   Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              75.8          Length:        60
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     2             Gaps:          0

US-09-606-910E-3 (1-7) x AAT88166 (1-60)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db          1 CCGGACACCAGGCCGGCCCCG 21

RESULT 6
AAV48320
ID   AAV48320 standard; DNA; 60 BP.
XX
AC   AAV48320;
XX
DT   20-NOV-1998  (first entry)
XX
DE   Nucleotide sequence encoding MUC1 tandem repeat unit R5.
XX
KW   ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW   tumour; tumour-associated antigen.
XX
OS   Homo sapiens.
XX
PN   WO9837095-A2.
XX
PD   27-AUG-1998.
XX
PF   24-FEB-1998;   98WO-US003693.
XX
PR   24-FEB-1997;   97US-0038253P.
XX
PA   (THER-) THERION BIOLOGICS CORP.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Schlom J,  Kantor J,  Kufe D,  Panicali D,  Gritz L;
XX
DR   WPI; 1998-467492/40.
XX
PT   New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT   immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS   Disclosure; Page 11; 42pp; English.
XX
CC   The MUC1 tandem repeat units AAV48317-V48325 were used to create an
```

```
CC   immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC   (RPV). The RPV was used in a pharmaceutical composition also containing
CC   an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC   recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC   does not undergo significant genetic deletion, thereby providing an
CC   unexpectedly stable and immunogenic pox virus. They can be used to
CC   prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ   Sequence 60 BP; 9 A; 27 C; 15 G; 9 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              75.8          Length:        60
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     2             Gaps:          0

US-09-606-910E-3 (1-7) x AAV48320 (1-60)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         40 CCGGATACCAGACCGGCCCCT 60

RESULT 7
AAV48321
ID   AAV48321 standard; DNA; 60 BP.
XX
AC   AAV48321;
XX
DT   20-NOV-1998  (first entry)
XX
DE   Nucleotide sequence encoding MUC1 tandem repeat unit R6.
XX
KW   ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW   tumour; tumour-associated antigen.
XX
OS   Homo sapiens.
XX
PN   WO9837095-A2.
XX
PD   27-AUG-1998.
XX
PF   24-FEB-1998;   98WO-US003693.
XX
PR   24-FEB-1997;   97US-0038253P.
XX
PA   (THER-) THERION BIOLOGICS CORP.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Schlom J,  Kantor J,  Kufe D,  Panicali D,  Gritz L;
XX
DR   WPI; 1998-467492/40.
XX
PT   New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT   immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS   Disclosure; Page 11; 42pp; English.
XX
CC   The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC   immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC   (RPV). The RPV was used in a pharmaceutical composition also containing
CC   an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC   recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC   does not undergo significant genetic deletion, thereby providing an
CC   unexpectedly stable and immunogenic pox virus. They can be used to
CC   prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ   Sequence 60 BP; 13 A; 27 C; 17 G; 3 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              75.8          Length:        60
```

```
Score:                     41.00          Matches:        7
Percent Similarity:        100.00%        Conservative:   0
Best Local Similarity:     100.00%        Mismatches:     0
Query Match:               100.00%        Indels:         0
DB:                        2              Gaps:           0

US-09-606-910E-3 (1-7) x AAV48321 (1-60)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         40 CCAGACACTCGACCTGCGCCA 60

RESULT 8
AAV48325
ID   AAV48325 standard; DNA; 60 BP.
XX
AC   AAV48325;
XX
DT   20-NOV-1998  (first entry)
XX
DE   Nucleotide sequence encoding MUC1 tandem repeat unit R10.
XX
KW   ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW   tumour; tumour-associated antigen.
XX
OS   Homo sapiens.
XX
PN   WO9837095-A2.
XX
PD   27-AUG-1998.
XX
PF   24-FEB-1998;   98WO-US003693.
XX
PR   24-FEB-1997;   97US-0038253P.
XX
PA   (THER-) THERION BIOLOGICS CORP.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Schlom J,  Kantor J,  Kufe D,  Panicali D,  Gritz L;
XX
DR   WPI; 1998-467492/40.
XX
PT   New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT   immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS   Disclosure; Page 11; 42pp; English.
XX
CC   The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC   immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC   (RPV). The RPV was used in a pharmaceutical composition also containing
CC   an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC   recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC   does not undergo significant genetic deletion, thereby providing an
CC   unexpectedly stable and immunogenic pox virus. They can be used to
CC   prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ   Sequence 60 BP; 12 A; 22 C; 19 G; 7 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                 75.8           Length:         60
Score:                     41.00          Matches:        7
Percent Similarity:        100.00%        Conservative:   0
Best Local Similarity:     100.00%        Mismatches:     0
Query Match:               100.00%        Indels:         0
DB:                        2              Gaps:           0

US-09-606-910E-3 (1-7) x AAV48325 (1-60)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         40 CCGGATACGAGACCGGCGCCT 60
```

```
RESULT 9
AAV48322
ID   AAV48322 standard; DNA; 60 BP.
XX
AC   AAV48322;
XX
DT   20-NOV-1998  (first entry)
XX
DE   Nucleotide sequence encoding MUC1 tandem repeat unit R7.
XX
KW   ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW   tumour; tumour-associated antigen.
XX
OS   Homo sapiens.
XX
PN   WO9837095-A2.
XX
PD   27-AUG-1998.
XX
PF   24-FEB-1998;   98WO-US003693.
XX
PR   24-FEB-1997;   97US-0038253P.
XX
PA   (THER-) THERION BIOLOGICS CORP.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Schlom J,  Kantor J,  Kufe D,  Panicali D,  Gritz L;
XX
DR   WPI; 1998-467492/40.
XX
PT   New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT   immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS   Disclosure; Page 11; 42pp; English.
XX
CC   The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC   immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC   (RPV). The RPV was used in a pharmaceutical composition also containing
CC   an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC   recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC   does not undergo significant genetic deletion, thereby providing an
CC   unexpectedly stable and immunogenic pox virus. They can be used to
CC   prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ   Sequence 60 BP; 14 A; 24 C; 16 G; 6 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                 75.8           Length:         60
Score:                     41.00          Matches:        7
Percent Similarity:        100.00%        Conservative:   0
Best Local Similarity:     100.00%        Mismatches:     0
Query Match:               100.00%        Indels:         0
DB:                        2              Gaps:           0

US-09-606-910E-3 (1-7) x AAV48322 (1-60)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         40 CCAGACACTCGACCTGCGCCA 60

RESULT 10
AAV48324
ID   AAV48324 standard; DNA; 60 BP.
XX
AC   AAV48324;
XX
DT   20-NOV-1998  (first entry)
XX
DE   Nucleotide sequence encoding MUC1 tandem repeat unit R9.
XX
KW   ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW   tumour; tumour-associated antigen.
```

```
XX
OS   Homo sapiens.
XX
PN   WO9837095-A2.
XX
PD   27-AUG-1998.
XX
PF   24-FEB-1998;   98WO-US003693.
XX
PR   24-FEB-1997;   97US-0038253P.
XX
PA   (THER-) THERION BIOLOGICS CORP.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Schlom J,  Kantor J,  Kufe D,  Panicali D,  Gritz L;
XX
DR   WPI; 1998-467492/40.
XX
PT   New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT   immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS   Disclosure; Page 11; 42pp; English.
XX
CC   The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC   immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC   (RPV). The RPV was used in a pharmaceutical composition also containing
CC   an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC   recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC   does not undergo significant genetic deletion, thereby providing an
CC   unexpectedly stable and immunogenic pox virus. They can be used to
CC   prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ   Sequence 60 BP; 12 A; 26 C; 14 G; 8 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              75.8          Length:        60
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     2             Gaps:          0

US-09-606-910E-3 (1-7) x AAV48324 (1-60)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         40 CCCGACACCCGTCCAGCTCCG 60

RESULT 11
AAV48318
ID   AAV48318 standard; DNA; 60 BP.
XX
AC   AAV48318;
XX
DT   20-NOV-1998  (first entry)
XX
DE   Nucleotide sequence encoding MUC1 tandem repeat unit R3.
XX
KW   ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW   tumour; tumour-associated antigen.
XX
OS   Homo sapiens.
XX
PN   WO9837095-A2.
XX
PD   27-AUG-1998.
XX
PF   24-FEB-1998;   98WO-US003693.
XX
PR   24-FEB-1997;   97US-0038253P.
XX
PA   (THER-) THERION BIOLOGICS CORP.
```

```
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Schlom J,  Kantor J,  Kufe D,  Panicali D,  Gritz L;
XX
DR   WPI; 1998-467492/40.
XX
PT   New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT   immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS   Disclosure; Page 11; 42pp; English.
XX
CC   The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC   immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC   (RPV). The RPV was used in a pharmaceutical composition also containing
CC   an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC   recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC   does not undergo significant genetic deletion, thereby providing an
CC   unexpectedly stable and immunogenic pox virus. They can be used to
CC   prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ   Sequence 60 BP; 7 A; 29 C; 19 G; 5 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              75.8          Length:        60
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     2             Gaps:          0

US-09-606-910E-3 (1-7) x AAV48318 (1-60)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         40 CCCGACACGCGCCCCGCTCCC 60

RESULT 12
AAV48323
ID   AAV48323 standard; DNA; 60 BP.
XX
AC   AAV48323;
XX
DT   20-NOV-1998  (first entry)
XX
DE   Nucleotide sequence encoding MUC1 tandem repeat unit R8.
XX
KW   ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW   tumour; tumour-associated antigen.
XX
OS   Homo sapiens.
XX
PN   WO9837095-A2.
XX
PD   27-AUG-1998.
XX
PF   24-FEB-1998;   98WO-US003693.
XX
PR   24-FEB-1997;   97US-0038253P.
XX
PA   (THER-) THERION BIOLOGICS CORP.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Schlom J,  Kantor J,  Kufe D,  Panicali D,  Gritz L;
XX
DR   WPI; 1998-467492/40.
XX
PT   New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT   immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS   Disclosure; Page 11; 42pp; English.
XX
```

```
CC   The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC   immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC   (RPV). The RPV was used in a pharmaceutical composition also containing
CC   an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC   recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC   does not undergo significant genetic deletion, thereby providing an
CC   unexpectedly stable and immunogenic pox virus. They can be used to
CC   prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ   Sequence 60 BP; 10 A; 24 C; 17 G; 9 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              75.8          Length:        60
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     2             Gaps:          0

US-09-606-910E-3 (1-7) x AAV48323 (1-60)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         40 CCTGACACAAGGCCAGCCCCA 60

RESULT 13
AAV48316
ID   AAV48316 standard; cDNA; 60 BP.
XX
AC   AAV48316;
XX
DT   20-NOV-1998  (first entry)
XX
DE   Nucleotide sequence encoding MUC1 tandem repeat unit.
XX
KW   ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW   tumour; tumour-associated antigen.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .60
FT                   /*tag=  a
FT                   /product= "MUC1 tandem repeat unit"
XX
PN   WO9837095-A2.
XX
PD   27-AUG-1998.
XX
PF   24-FEB-1998;   98WO-US003693.
XX
PR   24-FEB-1997;   97US-0038253P.
XX
PA   (THER-) THERION BIOLOGICS CORP.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Schlom J,  Kantor J,  Kufe D,  Panicali D,  Gritz L;
XX
DR   WPI; 1998-467492/40.
DR   P-PSDB; AAW77229.
XX
PT   New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT   immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS   Example 1; Page 20; 42pp; English.
XX
CC   The MUC1 tandem repeat unit was used to create an immunogenic mini-MUC1
CC   fragment for inclusion in a recombinant pox virus (RPV).. The RPV was used
CC   in a pharmaceutical composition also containing an immunomodulator to
CC   generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
CC   therefore encodes an immunogenic MUC1 fragment that does not undergo
CC   significant genetic deletion, thereby providing an unexpectedly stable
CC   and immunogenic pox virus. They can be used to prevent or treat tumours
CC   expressing MUC1 tumour-associated antigens
XX
SQ   Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              75.8          Length:        60
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     2             Gaps:          0

US-09-606-910E-3 (1-7) x AAV48316 (1-60)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         40 CCGGACACCAGGCCGGCCCCG 60

RESULT 14
AAV48319
ID   AAV48319 standard; DNA; 60 BP.
XX
AC   AAV48319;
XX
DT   20-NOV-1998  (first entry)
XX
DE   Nucleotide sequence encoding MUC1 tandem repeat unit R4.
XX
KW   ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW   tumour; tumour-associated antigen.
XX
OS   Homo sapiens.
XX
PN   WO9837095-A2.
XX
PD   27-AUG-1998.
XX
PF   24-FEB-1998;   98WO-US003693.
XX
PR   24-FEB-1997;   97US-0038253P.
XX
PA   (THER-) THERION BIOLOGICS CORP.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Schlom J,  Kantor J,  Kufe D,  Panicali D,  Gritz L;
XX
DR   WPI; 1998-467492/40.
XX
PT   New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT   immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS   Disclosure; Page 11; 42pp; English.
XX
CC   The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC   immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC   (RPV). The RPV was used in a pharmaceutical composition also containing
CC   an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC   recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC   does not undergo significant genetic deletion, thereby providing an
CC   unexpectedly stable and immunogenic pox virus. They can be used to
CC   prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ   Sequence 60 BP; 10 A; 23 C; 13 G; 14 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              75.8          Length:        60
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     2             Gaps:          0
```

```
US-09-606-910E-3 (1-7) x AAV48319 (1-60)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         40 CCAGATACTCGCCCAGCTCCA 60

RESULT 15
AAV48317
ID   AAV48317 standard; DNA; 60 BP.
XX
AC   AAV48317;
XX
DT   20-NOV-1998  (first entry)
XX
DE   Nucleotide sequence encoding MUC1 tandem repeat unit R2.
XX
KW   ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW   tumour; tumour-associated antigen.
XX
OS   Homo sapiens.
XX
PN   WO9837095-A2.
XX
PD   27-AUG-1998.
XX
PF   24-FEB-1998;   98WO-US003693.
XX
PR   24-FEB-1997;   97US-0038253P.
XX
PA   (THER-) THERION BIOLOGICS CORP.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Schlom J,  Kantor J,  Kufe D,  Panicali D,  Gritz L;
XX
DR   WPI; 1998-467492/40.
XX
PT   New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT   immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS   Disclosure; Page 11; 42pp; English.
XX
CC   The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC   immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC   (RPV). The RPV was used in a pharmaceutical composition also containing
CC   an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC   recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC   does not undergo significant genetic deletion, thereby providing an
CC   unexpectedly stable and immunogenic pox virus. They can be used to
CC   prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ   Sequence 60 BP; 17 A; 21 C; 13 G; 9 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              75.8          Length:        60
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     2             Gaps:          0

US-09-606-910E-3 (1-7) x AAV48317 (1-60)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         40 CCTGATACAAGACCTGCACCT 60


Search completed: December 29, 2004, 13:38:36
Job time : 334 secs
```

```
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Xing,Px. and McKenzie,I.F.C.
  TITLE     Monoclonal antibodies specific for human polymorphic epithelial
            mucin
  JOURNAL   Patent: EP 0369816-A 27 23-MAY-1990;
            THE UNIVERSITY OF MELBOURNE
FEATURES             Location/Qualifiers
     source          1..60
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:              70.7          Length:       60
Score:                  41.00         Matches:      7
Percent Similarity:     100.00%       Conservative: 0
Best Local Similarity:  100.00%       Mismatches:   0
Query Match:            100.00%       Indels:       0
DB:                     6             Gaps:         0

US-09-606-910E-3 (1-7) x A09845 (1-60)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db          1 CCGGACACCAGGCCGGCCCCG 21

RESULT 6
AX133974
LOCUS       AX133974                  60 bp    DNA     linear   PAT 15-MAY-2001
DEFINITION  Sequence 1 from Patent WO0124832.
ACCESSION   AX133974
VERSION     AX133974.1  GI:14139915
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Pecher,G.
  TITLE     Pharmaceutical composition for treating and preventing human
            tumors, which express the tumor antigen mucin and/or the
            carcinoembryonic antigen (ce a), and the use thereof
  JOURNAL   Patent: WO 0124832-A 1 12-APR-2001;
            Pecher, Gabriele (DE)
FEATURES             Location/Qualifiers
     source          1..60
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
     CDS             <1..>60
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC38972.1"
                     /db_xref="GI:14139916"
                     /translation="PDTRPAPGSTAPPAHGATSA"
ORIGIN

Alignment Scores:
Pred. No.:              70.7          Length:       60
Score:                  41.00         Matches:      7
Percent Similarity:     100.00%       Conservative: 0
Best Local Similarity:  100.00%       Mismatches:   0
Query Match:            100.00%       Indels:       0
DB:                     6             Gaps:         0

US-09-606-910E-3 (1-7) x AX133974 (1-60)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
```

```
Db          1 CCGGACACCAGGCCGGCCCCG 21

RESULT 7
HUMMUCIN
LOCUS       HUMMUCIN                  60 bp    mRNA    linear   PRI 27-JUL-1994
DEFINITION  Human epithelial mucin tandem repeat sequence of the partial cds.
ACCESSION   M26316
VERSION     M26316.1  GI:516622
KEYWORDS    mucin.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Xing,P.X., Tjandra,J.J., Reynolds,K., McLaughlin,P.J., Purcell,D.F.
            and McKenzie,I.F.
  TITLE     Reactivity of anti-human milk fat globule antibodies with synthetic
            peptides
  JOURNAL   J. Immunol. 142 (10), 3503-3509 (1989)
  MEDLINE   89235154
   PUBMED   2715633
COMMENT     On Jul 28, 1994 this sequence version replaced gi:341467.
            Original source text: Homo sapiens cDNA to mRNA.
FEATURES             Location/Qualifiers
     source          1..60
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             <1..>60
                     /note="epithelial mucin tandem repeat sequence"
                     /codon_start=1
                     /product="mucin"
                     /protein_id="AAA36336.1"
                     /db_xref="GI:516623"
                     /translation="PDTRPAPGSTAPPAHGVTSA"
ORIGIN

Alignment Scores:
Pred. No.:              70.7          Length:       60
Score:                  41.00         Matches:      7
Percent Similarity:     100.00%       Conservative: 0
Best Local Similarity:  100.00%       Mismatches:   0
Query Match:            100.00%       Indels:       0
DB:                     9             Gaps:         0

US-09-606-910E-3 (1-7) x HUMMUCIN (1-60)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db          1 CCGGACACCAGGCCGGCCCCG 21

RESULT 8
AX180273
LOCUS       AX180273                  86 bp    DNA     linear   PAT 06-AUG-2001
DEFINITION  Sequence 31 from Patent WO0146228.
ACCESSION   AX180273
VERSION     AX180273.1  GI:15132244
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Crowe,J.S. and Ellis,J.H.
  TITLE     Nucleic acid vaccination
  JOURNAL   Patent: WO 0146228-A 31 28-JUN-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1..86
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Oligonucleotide"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              97.7         Length:        86
Score:                  41.00        Matches:       7
Percent Similarity:     100.00%      Conservative:  0
Best Local Similarity:  100.00%      Mismatches:    0
Query Match:            100.00%      Indels:        0
DB:                     6            Gaps:          0

US-09-606-910E-3 (1-7) x AX180273 (1-86)

Qy          1 ProAspThrArgProAlaPro 7
                |||||||||||||||||||||
Db         33 CCGGACACCAGGCCGGCCCCG 53

RESULT 9
AX180274/c
LOCUS       AX180274                  92 bp    DNA     linear   PAT 06-AUG-2001
DEFINITION  Sequence 32 from Patent WO0146228.
ACCESSION   AX180274
VERSION     AX180274.1  GI:15132245
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Crowe,J.S. and Ellis,J.H.
  TITLE     Nucleic acid vaccination
  JOURNAL   Patent: WO 0146228-A 32 28-JUN-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .92
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Oligonucleotide"
ORIGIN

Alignment Scores:
Pred. No.:              104          Length:        92
Score:                  41.00        Matches:       7
Percent Similarity:     100.00%      Conservative:  0
Best Local Similarity:  100.00%      Mismatches:    0
Query Match:            100.00%      Indels:        0
DB:                     6            Gaps:          0

US-09-606-910E-3 (1-7) x AX180274 (1-92)

Qy          1 ProAspThrArgProAlaPro 7
                |||||||||||||||||||||
Db         60 CCGGACACCAGGCCGGCCCCG 40

RESULT 10
AX703425
LOCUS       AX703425                  93 bp    DNA     linear   PAT 03-APR-2003
DEFINITION  Sequence 3 from Patent WO02086505.
ACCESSION   AX703425
VERSION     AX703425.1  GI:29538416
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Benson,R.S.
  TITLE     Intracellular analysis
  JOURNAL   Patent: WO 02086505-A 3 31-OCT-2002;
            THE VICTORIA UNIVERSITY OF MANCHESTER (GB)
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Artificial epitope construct"
```

```
ORIGIN

Alignment Scores:
Pred. No.:              105          Length:        93
Score:                  41.00        Matches:       7
Percent Similarity:     100.00%      Conservative:  0
Best Local Similarity:  100.00%      Mismatches:    0
Query Match:            100.00%      Indels:        0
DB:                     6            Gaps:          0

US-09-606-910E-3 (1-7) x AX703425 (1-93)

Qy          1 ProAspThrArgProAlaPro 7
                |||||||||||||||||||||
Db         17 CCTGACACCAGACCTGCCCCT 37

RESULT 11
AX703426/c
LOCUS       AX703426                  93 bp    DNA     linear   PAT 03-APR-2003
DEFINITION  Sequence 4 from Patent WO02086505.
ACCESSION   AX703426
VERSION     AX703426.1  GI:29538417
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Benson,R.S.
  TITLE     Intracellular analysis
  JOURNAL   Patent: WO 02086505-A 4 31-OCT-2002;
            THE VICTORIA UNIVERSITY OF MANCHESTER (GB)
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Artificial epitope construct"
ORIGIN

Alignment Scores:
Pred. No.:              105          Length:        93
Score:                  41.00        Matches:       7
Percent Similarity:     100.00%      Conservative:  0
Best Local Similarity:  100.00%      Mismatches:    0
Query Match:            100.00%      Indels:        0
DB:                     6            Gaps:          0

US-09-606-910E-3 (1-7) x AX703426 (1-93)

Qy          1 ProAspThrArgProAlaPro 7
                |||||||||||||||||||||
Db         78 CCGGACACAAGGCCCGCCCCA 58

RESULT 12
AX192396/c
LOCUS       AX192396                 120 bp    DNA     linear   PAT 15-AUG-2001
DEFINITION  Sequence 2 from Patent EP1103623.
ACCESSION   AX192396
VERSION     AX192396.1  GI:15210363
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Taylor-Papadimitriou,J., Burchell,J. and Gendler,S.
  TITLE     Human mucin core protein: nucleic acid probes, peptide fragments
            and antibodies thereto, and uses thereof in diagnostic and
            therapeutic methods
  JOURNAL   Patent: EP 1103623-A 2 30-MAY-2001;
            IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .120
```

```
                  /organism="Homo sapiens"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              132           Length:        120
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     6             Gaps:          0

US-09-606-910E-3 (1-7) x AX192396 (1-120)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db        108 CCGGACACCAGGCCGGCCCCG 88

RESULT 13
AX192397
LOCUS       AX192397                 120 bp    DNA     linear   PAT 15-AUG-2001
DEFINITION  Sequence 3 from Patent EP1103623.
ACCESSION   AX192397
VERSION     AX192397.1  GI:15210364
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Taylor-Papadimitriou,J., Burchell,J. and Gendler,S.
  TITLE     Human mucin core protein: nucleic acid probes, peptide fragments
            and antibodies thereto, and uses thereof in diagnostic and
            therapeutic methods
  JOURNAL   Patent: EP 1103623-A 3 30-MAY-2001;
            IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1..120
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              132           Length:        120
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     6             Gaps:          0

US-09-606-910E-3 (1-7) x AX192397 (1-120)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         13 CCGGACACCAGGCCGGCCCCG 33

RESULT 14
BD000571/c
LOCUS       BD000571                 120 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Human polymorphic epithelial mucin core protein and nucleic acid
            encoding  the protein.
ACCESSION   BD000571
VERSION     BD000571.1  GI:18623684
KEYWORDS    JP 2000333675-A/2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 120)
  AUTHORS   Papadimitrov,J.T., Jendora,S. and Bachieru,J.
  TITLE     Human polymorphic epithelial mucin core protein and nucleic acid
```

```
            encoding the protein
  JOURNAL   Patent: JP 2000333675-A 2 05-DEC-2000;
            IMPERIAL CANCER RESEARCH TECHNOLOGY LTD
COMMENT     OS   Homo sapiens (human)
            PN   JP 2000333675-A/2
            PD   05-DEC-2000
            PF   26-APR-2000 JP 2000125724
            PR   07-JAN-1987 GB       8700269,07-JAN-1987 GB       8700279 PR
            22-APR-1987 US        041306,09-NOV-1987 GB       8726172 PI   JOYCE
            TAYLOR PAPADIMITROV,SANDRA JENDORA,JOY BACHIERU PC
            C12N15/02,A61K38/00,A61K39/395,A61K39/395,A61K49/00,A61P35/00, PC
            C07K14/47,
            PC   C07K16/44,C12N5/10,C12P21/08//(C12N15/02,C12R1:91),(C12N5/10,
            PC   C12R1:91),
            PC   C12N15/00,A61K37/02,C12N5/00,(C12N15/00,C12R1:91),(C12N5/00,
            PC   C12R1:91)
            CC
            FH   Key             Location/Qualifiers
            FT   source          1..120
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1..120
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              132           Length:        120
Score:                  41.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     6             Gaps:          0

US-09-606-910E-3 (1-7) x BD000571 (1-120)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db        108 CCGGACACCAGGCCGGCCCCG 88

RESULT 15
BD000572
LOCUS       BD000572                 120 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Human polymorphic epithelial mucin core protein and nucleic acid
            encoding  the protein.
ACCESSION   BD000572
VERSION     BD000572.1  GI:18623685
KEYWORDS    JP 2000333675-A/3.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 120)
  AUTHORS   Papadimitrov,J.T., Jendora,S. and Bachieru,J.
  TITLE     Human polymorphic epithelial mucin core protein and nucleic acid
            encoding the protein
  JOURNAL   Patent: JP 2000333675-A 3 05-DEC-2000;
            IMPERIAL CANCER RESEARCH TECHNOLOGY LTD
COMMENT     OS   Homo sapiens (human)
            PN   JP 2000333675-A/3
            PD   05-DEC-2000
            PF   26-APR-2000 JP 2000125724
            PR   07-JAN-1987 GB       8700269,07-JAN-1987 GB       8700279 PR
            22-APR-1987 US        041306,09-NOV-1987 GB       8726172 PI   JOYCE
            TAYLOR PAPADIMITROV,SANDRA JENDORA,JOY BACHIERU PC
            C12N15/02,A61K38/00,A61K39/395,A61K39/395,A61K49/00,A61P35/00, PC
            C07K14/47,
            PC   C07K16/44,C12N5/10,C12P21/08//(C12N15/02,C12R1:91),(C12N5/10,
            PC   C12R1:91),
            PC   C12N15/00,A61K37/02,C12N5/00,(C12N15/00,C12R1:91),(C12N5/00,
            PC   C12R1:91)
            CC
```

```
            FH   Key             Location/Qualifiers
            FT   source          1. .120
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .120
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              132            Length:        120
Score:                  41.00          Matches:       7
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     6              Gaps:          0

US-09-606-910E-3 (1-7) x BD000572 (1-120)

Qy          1 ProAspThrArgProAlaPro 7
              |||||||||||||||||||||
Db         13 CCGGACACCAGGCCGGCCCCG 33


Search completed: December 29, 2004, 14:21:05
Job time : 2544 secs
```

```
RT   "High density O-glycosylation on tandem repeat peptide from secretory
RT   MUC1 of T47D breast cancer cells.";
RL   J. Biol. Chem. 274:18165-18172(1999).
RN   [20]
RP   POLYMORPHISM WITHIN THE REPEAT.
RX   MEDLINE=21359366; PubMed=11350974; DOI=10.1074/jbc.M103187200;
RA   Engelmann K., Baldus S.E., Hanisch F.-G.;
RT   "Identification and topology of variant sequences within individual
RT   repeat domains of the human epithelial tumor mucin MUC1.";
RL   J. Biol. Chem. 276:27764-27769(2001).
RN   [21]
RP   CHARACTERIZATION OF ISOFORM Y, AND MUTAGENESIS OF ASP-1116.
RX   MEDLINE=99211485; PubMed=10197628;
RA   Baruch A., Hartmann M.-L., Yoeli M., Adereth Y., Greenstein S.,
RA   Stadler Y., Skornik Y., Zaretsky J., Smorodinsky N.I., Keydar I.,
RA   Wreschner D.H.;
RT   "The breast cancer-associated MUC1 gene generates both a receptor and
RT   its cognate binding protein.";
RL   Cancer Res. 59:1552-1561(1999).
RN   [22]
RP   PARTIAL SEQUENCE, AND CHARACTERIZATION OF CLEAVAGE SITE.
RX   MEDLINE=21240104; PubMed=11341784; DOI=10.1006/bbrc.2001.4775;
RA   Parry S., Silverman H.S., McDermott K., Willis A., Hollingsworth M.A.,
RA   Harris A.;
RT   "Identification of MUC1 proteolytic cleavage sites in vivo.";
RL   Biochem. Biophys. Res. Commun. 283:715-720(2001).
RN   [23]
RP   CHARACTERIZATION.
RX   MEDLINE=21836452; PubMed=11847293;
RA   Wreschner D.H., McGuckin M.A., Williams S.J., Baruch A., Yoeli M.,
RA   Ziv R., Okun L., Zaretsky J., Smorodinsky N., Keydar I., Neophytou P.,
RA   Stacey M., Lin H.-H., Gordon S.;
RT   "Generation of ligand-receptor alliances by 'SEA' module-mediated
RT   cleavage of membrane-associated mucin proteins.";
RL   Protein Sci. 11:698-706(2002).
RN   [24]
RP   PHOSPHORYLATION.

  Query Match             100.0%;  Score 41;  DB 1;  Length 1255;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db        142 PDTRPAP 148


RESULT 3
Q9V052
ID   Q9V052      PRELIMINARY;      PRT;   514 AA.
AC   Q9V052;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Major extracellular endo-1,4-betaglucanase (Cellulase).
GN   Name=celB-like; OrderedLocusNames=PYRAB09400; ORFNames=PAB0632;
OS   Pyrococcus abyssi.
OC   Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC   Pyrococcus.
OX   NCBI_TaxID=29292;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GE5 / Orsay;
RX   MEDLINE=22511545; PubMed=12622808;
RA   Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,
RA   Poch O., Prieur D., Querellou J., Ripp R., Thierry J.-C.,
RA   Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;
RT   "An integrated analysis of the genome of the hyperthermophilic
RT   archaeon Pyrococcus abyssi.";
RL   Mol. Microbiol. 47:1495-1512(2003).
DR   EMBL; AJ248285; CAB49854.1; -.
DR   PIR; E75142; E75142.
DR   HSSP; P54583; 1ECE.
DR   GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   InterPro; IPR001547; Glyco_hydro_5.
DR   Pfam; PF00150; Cellulase; 1.
DR   PROSITE; PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
KW   Complete proteome.
SQ   SEQUENCE   514 AA;  59980 MW;  CF80CAB1525337C5 CRC64;

  Query Match             92.7%;  Score 38;  DB 2;  Length 514;
  Best Local Similarity   85.7%;  Pred. No. 1.6e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||:|
Db        101 PDTRPSP 107


RESULT 4
O18977
ID   O18977      PRELIMINARY;      PRT;  4135 AA.
AC   O18977;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Tenascin-X.
GN   Name=TN-X;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RX   MEDLINE=97426436; PubMed=9278449;
RA   Elefteriou F., Exposito J.Y., Garrone R., Lethias C.;
RT   "Characterization of flexilin, the bovine tenascin-X.";
RL   J. Biol. Chem. 272:22866-22874(1997).
CC   -!- SIMILARITY: Contains 19 EGF-like domains.
DR   EMBL; Y11915; CAA72671.1; -.
DR   PIR; T42629; T42629.
DR   HSSP; P05107; 1L3Y.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002181; Fibrinogen_C.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR006210; IEGF.
DR   Pfam; PF00008; EGF; 11.
DR   Pfam; PF00147; Fibrinogen_C; 1.
DR   Pfam; PF00041; fn3; 30.
DR   SMART; SM00181; EGF; 9.
DR   SMART; SM00186; FBG; 1.
DR   SMART; SM00060; FN3; 30.
DR   PROSITE; PS00022; EGF_1; 19.
DR   PROSITE; PS01186; EGF_2; 19.
DR   PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
DR   PROSITE; PS50853; FN3; 30.
KW   EGF-like domain.
SQ   SEQUENCE   4135 AA;  447381 MW;  D903653C3444F4A9 CRC64;

  Query Match             92.7%;  Score 38;  DB 2;  Length 4135;
  Best Local Similarity   85.7%;  Pred. No. 1.4e+03;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||:|||
Db       3240 PDTKPAP 3246


RESULT 5
PCLO_CHICK
ID   PCLO_CHICK     STANDARD;      PRT;  5120 AA.
AC   Q9PU36;
```

```
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Piccolo protein (Aczonin) (Fragment).
GN   Name=PCLO; Synonyms=ACZ;
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=99439764; PubMed=10508862;
RA   Wang X., Kibschull M., Laue M.M., Lichte B., Petrasch-Parwez E.,
RA   Kilimann M.W.;
RT   "Aczonin, a 550-kd putative scaffolding protein of presynaptic active
RT   zones, shares homology regions with rim and bassoon and binds
RT   profilin.";
RL   J. Cell Biol. 147:151-162(1999).
CC   -!- FUNCTION: May act as a scaffolding protein involved in the
CC       organization of synaptic active zones and in synaptic vesicle
CC       trafficking (By similarity).
CC   -!- SUBUNIT: Interacts with Rabac1/Pra1 and profilin (By similarity).
CC   -!- SUBCELLULAR LOCATION: Concentrated at the presynaptic side of
CC       synaptic junctions (By similarity).
CC   -!- DOMAIN: C2 domain 1 is involved in binding calcium and
CC       phospholipids. Calcium binds with low affinity but with high
CC       specificity and induces a large conformational change.
CC   -!- SIMILARITY: Contains 2 C2 domains.
CC   -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y19187; CAB60725.1; -.
DR   HSSP; P04410; 1A25.
DR   GO; GO:0045202; C:synapse; ISS.
DR   GO; GO:0005509; F:calcium ion binding; ISS.
DR   GO; GO:0005544; F:calcium-dependent phospholipid binding; ISS.
DR   GO; GO:0005522; F:profilin binding; ISS.
DR   GO; GO:0007010; P:cytoskeleton organization and biogenesis; ISS.
DR   GO; GO:0016080; P:synaptic vesicle targeting; ISS.
DR   InterPro; IPR000008; C2.
DR   InterPro; IPR001478; PDZ.
DR   InterPro; IPR001565; Synaptotagmin.
DR   InterPro; IPR008899; Znf_piccolo.
DR   Pfam; PF00168; C2; 2.
DR   Pfam; PF00595; PDZ; 1.
DR   Pfam; PF05715; zf-piccolo; 2.
DR   PRINTS; PR00399; SYNAPTOTAGMN.
DR   SMART; SM00239; C2; 2.
DR   SMART; SM00228; PDZ; 1.
DR   PROSITE; PS00499; C2_DOMAIN_1; 1.
DR   PROSITE; PS50004; C2_DOMAIN_2; 2.
DR   PROSITE; PS50106; PDZ; 1.
KW   Calcium/phospholipid-binding; Metal-binding; Repeat; Zinc;
KW   Zinc-finger.
FT   NON_TER       1      1
FT   DOMAIN      258    357       10 X 10 AA tandem approximate repeats of
FT                                P-A-K-P-Q-P-Q-Q-P-X.
FT   ZN_FING     368    392       C4-type (Potential).
FT   ZN_FING     836    859       C4-type (Potential).
FT   DOMAIN     2324   2343       Poly-Pro.
FT   DOMAIN     4414   4493       PDZ.
FT   DOMAIN     4627   4726       C2 domain 1.
FT   DOMAIN     5003   5094       C2 domain 2.
SQ   SEQUENCE   5120 AA;  560751 MW;  A658D9891B65B412 CRC64;
```

```
  Query Match             92.7%;  Score 38;  DB 1;  Length 5120;
  Best Local Similarity   85.7%;  Pred. No. 1.8e+03;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||:|||
Db        643 PDTKPAP 649


RESULT 6
Q84559
ID   Q84559      PRELIMINARY;      PRT;   111 AA.
AC   Q84559;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   A239L protein.
GN   Name=A239L;
OS   Paramecium bursaria chlorella virus 1 (PBCV-1).
OC   Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
OX   NCBI_TaxID=10506;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96187795; PubMed=8614977;
RA   Lu Z., Li Y., Que Q., Kutish G.F., Rock D.L., Van Etten J.L.;
RT   "Analysis of 94 kb of the chlorella virus PBCV-1 330-kb genome: map
RT   positions 88 to 182.";
RL   Virology 216:102-123(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20013326; PubMed=10544099;
RA   Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,
RA   Lisec A.D., Nickerson K.W., Van Etten J.L.;
RT   "Chlorella virus PBCV-1 encodes a functional homospermidine
RT   synthase.";
RL   Virology 263:254-262(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20478054; PubMed=11021991;
RA   Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;
RT   "Characterization of a beta-1,3-glucanase encoded by chlorella virus
RT   PBCV-1.";
RL   Virology 276:27-36(2000).
DR   EMBL; U42580; AAC96607.1; -.
DR   PIR; T17730; T17730.
SQ   SEQUENCE   111 AA;  12721 MW;  1CAC3DFF047017E8 CRC64;

  Query Match             90.2%;  Score 37;  DB 2;  Length 111;
  Best Local Similarity   85.7%;  Pred. No. 50;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              ||||| |
Db         13 PDTRPVP 19


RESULT 7
Q6J2G2
ID   Q6J2G2      PRELIMINARY;      PRT;   186 AA.
AC   Q6J2G2;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Single-stranded binding protein.
GN   ORFNames=PMA4326A42;
OS   Pseudomonas syringae (pv. maculicola).
OG   Plasmid pPMA4326A.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=59511;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=ES4326;
RX   PubMed=15262947;
RA   Stavrinides J., Guttman D.S.;
RT   "Nucleotide Sequence and Evolution of the Five-Plasmid Complement of
RT   the Phytopathogen Pseudomonas syringae pv. maculicola ES4326.";
RL   J. Bacteriol. 186:5101-5115(2004).
CC   -!- SIMILARITY: Contains 1 SSB domain.
DR   EMBL; AY603979; AAT35159.1; -.
DR   InterPro; IPR008994; Nucleic_acid_OB.
DR   InterPro; IPR011344; SSB.
DR   InterPro; IPR000424; SSB/primosomal_n.
DR   InterPro; IPR010913; SS_binding.
DR   Pfam; PF00436; SSB; 1.
DR   TIGRFAMs; TIGR00621; ssb; 1.
DR   PROSITE; PS50935; SSB; 1.
KW   DNA replication; DNA-binding; Plasmid.
SQ   SEQUENCE   186 AA;  20608 MW;  DF790B4F6FB20B2A CRC64;

  Query Match             90.2%;  Score 37;  DB 2;  Length 186;
  Best Local Similarity   85.7%;  Pred. No. 86;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              ||:||||
Db        160 PDSRPAP 166


RESULT 8
AAT35159
ID   AAT35159     PRELIMINARY;      PRT;   186 AA.
AC   AAT35159;
DT   01-JUN-2004 (TrEMBLrel. 27, Created)
DT   01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT   01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE   Single-stranded binding protein.
GN   PMA4326A42.
OS   Pseudomonas syringae (pv. maculicola).
OG   Plasmid pPMA4326A.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas; Pseudomonas syringae group genomosp. 3.
OX   NCBI_TaxID=59511;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ES4326;
RA   Stavrinides J., Guttman D.S.;
RT   "Nucleotide Sequence and Evolution of the Five-Plasmid Complement of
RT   the Phytopathogen Pseudomonas syringae pv. maculicola ES4326.";
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY603979; AAT35159.1; -.
KW   Plasmid.
SQ   SEQUENCE   186 AA;  20608 MW;  DF790B4F6FB20B2A CRC64;

  Query Match             90.2%;  Score 37;  DB 2;  Length 186;
  Best Local Similarity   85.7%;  Pred. No. 86;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              ||:||||
Db        160 PDSRPAP 166


RESULT 9
Q8QZS5
ID   Q8QZS5       PRELIMINARY;      PRT;   190 AA.
AC   Q8QZS5;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   RIKEN cDNA 5830442J12 (Mus musculus adult male thymus cDNA, RIKEN
DE   full-length enriched library, clone:5830442J12 product:hypothetical
DE   protein, full insert sequence).
GN   Name=5830442J12Rik;
OS   Mus musculus (Mouse).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Mix FVB/N;
RC   TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Mix FVB/N;
RC   TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Thymus;
RX   MEDLINE=99279253; PubMed=10349636;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Thymus;
RX   MEDLINE=21085660; PubMed=11217851;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Thymus;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Thymus;
RX   MEDLINE=20499374; PubMed=11042159;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [7]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Thymus;
RX   MEDLINE=20530913; PubMed=11076861;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
```

```
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [8]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Thymus;
RA   Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA   Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA   Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA   Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA   Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA   Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA   Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA   Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA   Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA   Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA   Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC025512; AAH25512.1; -.
DR   EMBL; AK017982; BAC25537.1; -.
DR   MGD; MGI:1923322; 5830442J12Rik.
DR   GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR   GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0016567; P:protein ubiquitination; IEA.
DR   InterPro; IPR001841; Znf_ring.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS00518; ZF_RING_1; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ   SEQUENCE   190 AA;  21625 MW;  0671BD5D56DB7FB1 CRC64;

  Query Match             90.2%;  Score 37;  DB 2;  Length 190;
  Best Local Similarity   85.7%;  Pred. No. 88;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              ||||| |
Db        136 PDTRPVP 142

RESULT 10
Q7QF20
ID   Q7QF20      PRELIMINARY;      PRT;   249 AA.
AC   Q7QF20;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   AgCP13330 (Fragment).
GN   Name=agCG54728; ORFNames=ENSANGG00000009896;
OS   Anopheles gambiae str. PEST.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX   NCBI_TaxID=180454;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PEST;
RA   Anopheles Genome Sequencing Consortium;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAAB01008846; EAA06446.1; -.
DR   GO; GO:0005489; F:electron transporter activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR002109; Glutaredoxin.
DR   Pfam; PF00462; Glutaredoxin; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   249 AA;  28152 MW;  F0610FB5E3379629 CRC64;
```

```
  Query Match             90.2%;  Score 37;  DB 2;  Length 249;
  Best Local Similarity   85.7%;  Pred. No. 1.2e+02;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              ||||| |
Db         28 PDTRPGP 34

RESULT 11
Q82QQ1
ID   Q82QQ1      PRELIMINARY;      PRT;   384 AA.
AC   Q82QQ1;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=SAV454;
OS   Streptomyces avermitilis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=33903;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680;
RX   MEDLINE=21477403; PubMed=11572948;
RA   Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA   Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA   Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT   "Genome sequence of an industrial microorganism Streptomyces
RT   avermitilis: deducing the ability of producing secondary
RT   metabolites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680;
RX   MEDLINE=22608306; PubMed=12692562;
RA   Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA   Sakaki Y., Hattori M., Omura S.;
RT   "Complete genome sequence and comparative analysis of the industrial
RT   microorganism Streptomyces avermitilis.";
RL   Nat. Biotechnol. 21:526-531(2003).
DR   EMBL; AP005023; BAC68164.1; -.
DR   InterPro; IPR000169; Pept_cys_acsite.
DR   PROSITE; PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   384 AA;  42158 MW;  2AB4A385EAD71046 CRC64;

  Query Match             90.2%;  Score 37;  DB 2;  Length 384;
  Best Local Similarity   85.7%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              ||||| |
Db        163 PDTRPGP 169

RESULT 12
Q88LE1
ID   Q88LE1      PRELIMINARY;      PRT;   911 AA.
AC   Q88LE1;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=PP1993;
OS   Pseudomonas putida (strain KT2440).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=160488;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22423060; PubMed=12534463;
```

```
RA   Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA   Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA   Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA   Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA   Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA   Moazzez A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA   Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA   Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA   Fraser C.M.;
RT   "Complete genome sequence and comparative analysis of the
RT   metabolically versatile Pseudomonas putida KT2440.";
RL   Environ. Microbiol. 4:799-808(2002).
DR   EMBL; AE016781; AAN67608.1; -.
DR   TIGR; PP1993; -.
DR   GO; GO:0016998; P:cell wall catabolism; IEA.
DR   InterPro; IPR003439; ABC_transporter.
DR   InterPro; IPR002482; LysM.
DR   InterPro; IPR008941; TPR-like.
DR   Pfam; PF01476; LysM; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   911 AA;  96918 MW;  F6EC7861AECFA2F4 CRC64;

  Query Match             90.2%;  Score 37;  DB 2;  Length 911;
  Best Local Similarity   85.7%;  Pred. No. 4.5e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||:|||
Db        366 PDTQPAP 372


RESULT 13
O15870
ID   O15870      PRELIMINARY;      PRT;  2706 AA.
AC   O15870;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   PfEMP1 (Fragment).
GN   Name=R29R+var1;
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IT 4/25/5;
RX   MEDLINE=97373957; PubMed=9230440;
RA   Rowe J.A., Moulds J.M., Newbold C.I., Miller L.H.;
RT   "Plasmodium falciparum rosetting is mediated by PfEMP1 and requires
RT   complement receptor 1.";
RL   Nature 388:292-295(1997).
DR   EMBL; Y13402; CAA73831.1; -.
DR   EMBL; Y13403; CAA73831.1; JOINED.
DR   PIR; T28155; T28155.
DR   GO; GO:0005539; F:glycosaminoglycan binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   InterPro; IPR002086; Aldehyde_dehydr.
DR   InterPro; IPR004258; PFEMP.
DR   Pfam; PF03011; PFEMP; 2.
DR   PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
FT   NON_TER    2706   2706
SQ   SEQUENCE   2706 AA;  308164 MW;  1C33D55AD5317D68 CRC64;

  Query Match             90.2%;  Score 37;  DB 2;  Length 2706;
  Best Local Similarity   85.7%;  Pred. No. 1.4e+03;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              ||||| |
Db       2450 PDTRPVP 2456
```

```
RESULT 14
Q84187
ID   Q84187      PRELIMINARY;      PRT;   125 AA.
AC   Q84187;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86200383; PubMed=3754588;
RA   Bellini W.J., Englund G., Richardson C.D., Rozenblatt S.,
RA   Lazzarini R.A.;
RT   "Matrix genes of measles virus and canine distemper virus: cloning,
RT   nucleotide sequences, and deduced amino acid sequences.";
RL   J. Virol. 58:408-416(1986).
DR   EMBL; M12668; AAA66617.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   125 AA;  13356 MW;  4C9482BA3D7A06F6 CRC64;

  Query Match             87.8%;  Score 36;  DB 2;  Length 125;
  Best Local Similarity   85.7%;  Pred. No. 86;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              ||||| |
Db         49 PDTRPPP 55


RESULT 15
Q8PLY4
ID   Q8PLY4      PRELIMINARY;      PRT;   198 AA.
AC   Q8PLY4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   TonB-like protein.
GN   OrderedLocusNames=XAC1651;
OS   Xanthomonas axonopodis (pv. citri).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=92829;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=306 / ATCC 13902 / XV 101;
RX   MEDLINE=22022145; PubMed=12024217;
RA   da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA   Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA   Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA   Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA   Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA   Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA   Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA   Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA   Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA   Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA   Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA   Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA   Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA   Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA   Setubal J.C., Kitajima J.P.;
RT   "Comparison of the genomes of two Xanthomonas pathogens with differing
RT   host specificities.";
RL   Nature 417:459-463(2002).
DR   EMBL; AE011797; AAM36519.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR   GO; GO:0008565; F:protein transporter activity; IEA.
DR   GO; GO:0015031; P:protein transport; IEA.
```

```
DR   InterPro; IPR006260; TonB_C.
DR   TIGRFAMs; TIGR01352; tonB_Cterm; 1.
KW   Complete proteome.
SQ   SEQUENCE   198 AA;  21748 MW;  FB9DF173EDC52961 CRC64;

  Query Match             87.8%;  Score 36;  DB 2;  Length 198;
  Best Local Similarity   85.7%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
                ||||| |
Db         70 PDTRPLP 76


Search completed: December 29, 2004, 13:11:13
Job time : 97 secs
```

```
Db        101 PDTRPSP 107


RESULT 4
T42629
tenascin-X - bovine
N;Alternate names: flexilin
C;Species: Bos primigenius taurus (cattle)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42629
R;Elefteriou, F.; Exposito, J.Y.; Garrone, R.; Lethias, C.
J. Biol. Chem. 272, 22866-22874, 1997
A;Title: Characterization of the bovine tenascin-X.
A;Reference number: Z22180; MUID:97426436; PMID:9278449
A;Accession: T42629
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-4135 <ELE>
A;Cross-references: UNIPROT:O18977; EMBL:Y11915; NID:g2462978; PIDN:CAA72671.1; PID:g246
C;Genetics:
A;Gene: TN-X
C;Superfamily: tenascin-X; EGF homology; fibrinogen beta/gamma homology; fibronectin typ
C;Keywords: extracellular matrix; glycoprotein; heptad repeat

  Query Match             92.7%;  Score 38;  DB 2;  Length 4135;
  Best Local Similarity   85.7%;  Pred. No. 1.7e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||:|||
Db       3240 PDTKPAP 3246


RESULT 5
T17730
hypothetical protein A239L - Chlorella virus PBCV-1
C;Species: Chlorella virus PBCV-1
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17730
R;Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z18806
A;Accession: T17730
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-111 <GRA>
A;Cross-references: UNIPROT:Q84559; EMBL:U42580; NID:g4028896; PIDN:AAC96607.1
A;Experimental source: specific host Chlorella strain NC64A
C;Genetics:
A;Note: A239L
C;Superfamily: Chlorella virus PBCV-1 hypothetical protein A239L

  Query Match             90.2%;  Score 37;  DB 2;  Length 111;
  Best Local Similarity   85.7%;  Pred. No. 6.9;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              ||||| |
Db         13 PDTRPVP 19


RESULT 6
T28155
variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum) (fragments
N;Alternate names: erythrocyte membrane binding protein 1 (EMP1)
C;Species: Plasmodium falciparum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T28155
R;Rowe, J.A.; Moulds, J.M.; Newbold, C.I.; Miller, L.H.
Nature 388, 292-295, 1997
A;Title: Plasmodium falciparum rosetting is mediated by PfEMP1 and requires complement r
A;Reference number: Z20477; MUID:97373957; PMID:9230440
A;Accession: T28155
```

```
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2706 <ROW>
A;Cross-references: UNIPROT:O15870; EMBL:Y13402; PIDN:CAA73831.1
A;Experimental source: strain IT 4/25/5
C;Genetics:
A;Introns: 2493/3
A;Note: R29R+var1

  Query Match             90.2%;  Score 37;  DB 2;  Length 2706;
  Best Local Similarity   85.7%;  Pred. No. 1.7e+02;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              ||||| |
Db       2450 PDTRPVP 2456


RESULT 7
G82129
hypothetical protein VC2005 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: G82129
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: G82129
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-118 <HEI>
A;Cross-references: UNIPROT:Q9KQJ3; GB:AE004275; GB:AE003852; NID:g9656544; PIDN:AAF9515
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC2005
A;Map position: 1

  Query Match             85.4%;  Score 35;  DB 2;  Length 118;
  Best Local Similarity   85.7%;  Pred. No. 17;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              ||| |||
Db         26 PDTAPAP 32


RESULT 8
D83142
hypothetical protein PA4040 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: D83142
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathc
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: D83142
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-335 <STO>
A;Cross-references: UNIPROT:Q9HWY7; GB:AE004820; GB:AE004091; NID:g9950223; PIDN:AAG0742
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA4040

  Query Match             85.4%;  Score 35;  DB 2;  Length 335;
  Best Local Similarity   85.7%;  Pred. No. 48;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          1 PDTRPAP 7
              ||| |||
Db        308 PDTAPAP 314


RESULT 9
AI3083
monooxygenase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AI3083
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AI3083
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-341 <KUR>
A;Cross-references: UNIPROT:Q8U806; GB:AE008689; PIDN:AAL45087.1; PID:g17742754; GSPDB:G
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu4293
A;Map position: linear chromosome
C;Superfamily: alkanal monooxygenase [FMN-linked] (bacterial luciferase)

  Query Match             85.4%;  Score 35;  DB 2;  Length 341;
  Best Local Similarity   85.7%;  Pred. No. 49;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              || ||||
Db        270 PDERPAP 276


RESULT 10
H98202
hypothetical protein AGR_L_1143 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: H98202
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: H98202
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-355 <KUR>
A;Cross-references: UNIPROT:Q8U806; GB:AE007870; PIDN:AAK89146.1; PID:g15158956; GSPDB:G
C;Genetics:
A;Gene: AGR_L_1143
A;Map position: linear chromosome

  Query Match             85.4%;  Score 35;  DB 2;  Length 355;
  Best Local Similarity   85.7%;  Pred. No. 51;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              || ||||
Db        284 PDERPAP 290


RESULT 11
E81838
hypothetical protein NMA1474 [imported] - Neisseria meningitidis (strain Z2491 serogroup
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: E81838
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: E81838
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-365 <PAR>
A;Cross-references: UNIPROT:Q9JU64; GB:AL162756; GB:AL157959; NID:g7380091; PIDN:CAB847C
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA1474
C;Superfamily: Neisseria meningitidis hypothetical protein NMB1269

  Query Match             85.4%;  Score 35;  DB 2;  Length 365;
  Best Local Similarity   85.7%;  Pred. No. 52;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              ||||| |
Db          8 PDTRPYP 14


RESULT 12
G81102
hypothetical protein NMB1269 [imported] - Neisseria meningitidis (strain MC58 serogroup
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: G81102
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: G81102
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-365 <TET>
A;Cross-references: UNIPROT:Q9JZ68; GB:AE002475; GB:AE002098; NID:g7226502; PIDN:AAF4164
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB1269
C;Superfamily: Neisseria meningitidis hypothetical protein NMB1269

  Query Match             85.4%;  Score 35;  DB 2;  Length 365;
  Best Local Similarity   85.7%;  Pred. No. 52;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              ||||| |
Db          8 PDTRPYP 14


RESULT 13
G84302
hypothetical protein Vng1484h [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: G84302
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
 Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: G84302
A;Status: preliminary
```

```
A;Molecule type: DNA
A;Residues: 1-422 <STO>
A;Cross-references: UNIPROT:Q9HPT0; GB:AE004437; NID:g10580983; PIDN:AAG19787.1; GSPDB:G
C;Genetics:
A;Gene: VNG1484H

  Query Match             85.4%;  Score 35;  DB 2;  Length 422;
  Best Local Similarity   85.7%;  Pred. No. 61;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              || ||||
Db        176 PDRRPAP 182


RESULT 14
T05839
subtilisin-like proteinase homolog F17L22.100 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05839
R;Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Ba
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15454
A;Accession: T05839
A;Molecule type: DNA
A;Residues: 1-769 <BEV>
A;Cross-references: UNIPROT:Q9SVT3; EMBL:AL035527; GSPDB:GN00062; ATSP:F17L22.100
A;Experimental source: cultivar Columbia; BAC clone F17L22
C;Genetics:
A;Gene: ATSP:F17L22.100
A;Map position: 4
A;Introns: 20/1; 79/1; 141/1; 240/1; 363/3; 458/2; 496/3
C;Superfamily: subtilisin-like proteinase ag12; subtilisin homology

  Query Match             85.4%;  Score 35;  DB 2;  Length 769;
  Best Local Similarity   85.7%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              || ||||
Db        429 PDDRPAP 435


RESULT 15
A43359
microtubule-associated protein MAP1A - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: A43359; S22108
R;Langkopf, A.; Hammarback, J.A.; Muller, R.; Vallee, R.B.; Garner, C.C.
J. Biol. Chem. 267, 16561-16566, 1992
A;Title: Microtubule-associated proteins 1A and LC2. Two proteins encoded in one messeng
A;Reference number: A43359; MUID:92355629; PMID:1379599
A;Accession: A43359
A;Molecule type: mRNA
A;Residues: 1-2774 <LAN>
A;Cross-references: UNIPROT:P34926; GB:M83196; NID:g205537; PIDN:AAB48069.1; PID:g205538
A;Note: sequence extracted from NCBI backbone (NCBIN:111039, NCBIP:111040)
R;Cravchik, A.
submitted to the EMBL Data Library, June 1992
A;Reference number: S22108
A;Accession: S22108
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 73-364,'NRLRS',370,'QKN',374,'PSPKGL',381-751,'RSMMSQMNAQRR',764,'D',766,'LR
'WLKRNMCPQPRQSP',851,'V',853,'NSL',855,'LPHRWLRTN',865,'W',867,'HSQLPDGGD',877,'Q',879,'
A;Cross-references: EMBL:X66840
A;Experimental source: strain Sprague Dawley
C;Superfamily: microtubule-associated protein MAP1B
C;Keywords: microtubule binding; phosphoprotein

  Query Match             85.4%;  Score 35;  DB 2;  Length 2774;
```

```
  Best Local Similarity   85.7%;  Pred. No. 4e+02;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              || ||||
Db        757 PDERPAP 763

Search completed: December 29, 2004, 13:13:00
Job time : 26 secs
```

```
;  PRIOR FILING DATE: 1999-01-25
;  PRIOR APPLICATION NUMBER: 08/542,003
;  PRIOR FILING DATE: 1995-10-12
;  PRIOR APPLICATION NUMBER: 08/322,760
;  PRIOR FILING DATE: 1994-10-13
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: artificial
;   FEATURE:
;   OTHER INFORMATION: Antibody Sm3 Complex With Its Peptide Epitope
US-10-371-067-7

  Query Match             100.0%;  Score 41;  DB 15;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 4.1;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          4 PDTRPAP 10


RESULT 7
US-09-822-698A-40
; Sequence 40, Application US/09822698A
; Patent No. US20020146750A1
; GENERAL INFORMATION:
;  APPLICANT: Hoogenboom, Hendricus R.J.M.
;  APPLICANT:  Henderikx, Maria P.G.
;  TITLE OF INVENTION: MUCIN-1 Specific Binding Members and Methods of Use Thereof
;  FILE REFERENCE: DYX-015.1 US
;  CURRENT APPLICATION NUMBER: US/09/822,698A
;  CURRENT FILING DATE:  2001-03-30
;  PRIOR APPLICATION NUMBER: US 09/538,913
;  PRIOR FILING DATE: 2000-03-30
;  NUMBER OF SEQ ID NOS: 112
;  SOFTWARE: Microsoft Word
; SEQ ID NO 40
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: PAP peptide epitope of MUC1 bound by PH1 Fab antibody
US-09-822-698A-40

  Query Match             100.0%;  Score 41;  DB 9;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 4.7;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          1 PDTRPAP 7


RESULT 8
US-09-984-183-17
; Sequence 17, Application US/09984183
; Patent No. US20020142983A1
; GENERAL INFORMATION:
;  APPLICANT: AGRAWAL, BABITA
;  APPLICANT:  LONGENECKER, MICHAEL B.
;  TITLE OF INVENTION: MUC-1 ANTAGONISTS AND METHODS OF TREATING IMMUNE
;  TITLE OF INVENTION:  DISORDERS
;  FILE REFERENCE: 042881/0130
;  CURRENT APPLICATION NUMBER: US/09/984,183
;  CURRENT FILING DATE:  2001-10-29
;  PRIOR APPLICATION NUMBER: 09/457,354
;  PRIOR FILING DATE: 2001-10-29
;  PRIOR APPLICATION NUMBER: 60/111,973
;  PRIOR FILING DATE: 1998-12-11
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-984-183-17

  Query Match             100.0%;  Score 41;  DB 9;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          6 PDTRPAP 12


RESULT 9
US-09-984-333-7
; Sequence 7, Application US/09984333
; Patent No. US20020159969A1
; GENERAL INFORMATION:
;  APPLICANT: AGRAWAL, Babita
;  APPLICANT:  LONGENECKER, Bryan Michael
;  APPLICANT:  REDDISH, Mark Austin
;  TITLE OF INVENTION: SMALL PEPTIDE-BASED THERAPEUTICS FOR REVERSING
;  TITLE OF INVENTION:  CANCER-ASSOCIATED MUC-1 MUCIN-INDUCED IMMUNOSUPPRESSION
;  FILE REFERENCE: 042881/0116
;  CURRENT APPLICATION NUMBER: US/09/984,333
;  CURRENT FILING DATE:  2001-10-29
;  PRIOR APPLICATION NUMBER: US 09/182,887
;  PRIOR FILING DATE: 1998-10-30
;  PRIOR APPLICATION NUMBER: 60/064,146
;  PRIOR FILING DATE: 1997-10-31
;  PRIOR APPLICATION NUMBER: 60/065,209
;  PRIOR FILING DATE: 1997-11-12
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-984-333-7

  Query Match             100.0%;  Score 41;  DB 9;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          6 PDTRPAP 12


RESULT 10
US-09-143-379-1
; Sequence 1, Application US/09143379
; Publication No. US20040077826A1
; GENERAL INFORMATION:
;  APPLICANT: KOGANTY, R. Rao
;  APPLICANT: QIU, Dongxu
;  APPLICANT: GANDHI, Sham
;  TITLE OF INVENTION: RANDOMLY GENERATED GLYCOPEPTIDE COMBINATORIAL LIBRARIES
;  FILE REFERENCE: 042881/0119
;  CURRENT APPLICATION NUMBER: US/09/143,379
;  CURRENT FILING DATE: 1998-08-28
;  EARLIER APPLICATION NUMBER: 60/056,240
;  EARLIER FILING DATE: 1997-08-28
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
```

```
;   OTHER INFORMATION: Description of Artificial Sequence: Peptide
;   OTHER INFORMATION: derived from cancer-associated MUC1
US-09-143-379-1

  Query Match             100.0%;  Score 41;  DB 11;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          6 PDTRPAP 12

RESULT 11
US-09-815-346-3
; Sequence 3, Application US/09815346
; Publication No. US20020018806A1
; GENERAL INFORMATION:
;  APPLICANT: AGRAWAL, BABITA
;  APPLICANT:  LONGENECKER, MICHAEL B.
;  APPLICANT:  PARKER, JOANNE
;  TITLE OF INVENTION: LIPOPEPTIDE ADJUVANTS
;  FILE REFERENCE: 042881/0160
;  CURRENT APPLICATION NUMBER: US/09/815,346
;  CURRENT FILING DATE:  2001-03-23
;  PRIOR APPLICATION NUMBER: 60/191,736
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 18
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-09-815-346-3

  Query Match             100.0%;  Score 41;  DB 9;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 5.6;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          6 PDTRPAP 12

RESULT 12
US-10-106-876-6
; Sequence 6, Application US/10106876
; Publication No. US20030157160A1
; GENERAL INFORMATION:
;  APPLICANT: BUDZYNSKI, WLADYSLAW A.
;  APPLICANT:  KOGANTY, R. RAO
;  APPLICANT:  KRANTZ, MARK J.
;  APPLICANT:  LONGENECKER, B. MICHAEL
;  TITLE OF INVENTION: VACCINE FOR MODULATING BETWEEN T1 AND T2 IMMUNE
;  TITLE OF INVENTION:  RESPONSES
;  FILE REFERENCE: 042881-0176
;  CURRENT APPLICATION NUMBER: US/10/106,876
;  CURRENT FILING DATE:  2002-03-27
;  PRIOR APPLICATION NUMBER: 60/278,698
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 18
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-106-876-6

  Query Match             100.0%;  Score 41;  DB 14;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 5.6;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          6 PDTRPAP 12

RESULT 13
US-10-106-876-17
; Sequence 17, Application US/10106876
; Publication No. US20030157160A1
; GENERAL INFORMATION:
;  APPLICANT: BUDZYNSKI, WLADYSLAW A.
;  APPLICANT:  KOGANTY, R. RAO
;  APPLICANT:  KRANTZ, MARK J.
;  APPLICANT:  LONGENECKER, B. MICHAEL
;  TITLE OF INVENTION: VACCINE FOR MODULATING BETWEEN T1 AND T2 IMMUNE
;  TITLE OF INVENTION:  RESPONSES
;  FILE REFERENCE: 042881-0176
;  CURRENT APPLICATION NUMBER: US/10/106,876
;  CURRENT FILING DATE:  2002-03-27
;  PRIOR APPLICATION NUMBER: 60/278,698
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
;   LENGTH: 18
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-106-876-17

  Query Match             100.0%;  Score 41;  DB 14;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 5.6;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          7 PDTRPAP 13

RESULT 14
US-09-815-346-2
; Sequence 2, Application US/09815346
; Publication No. US20020018806A1
; GENERAL INFORMATION:
;  APPLICANT: AGRAWAL, BABITA
;  APPLICANT:  LONGENECKER, MICHAEL B.
;  APPLICANT:  PARKER, JOANNE
;  TITLE OF INVENTION: LIPOPEPTIDE ADJUVANTS
;  FILE REFERENCE: 042881/0160
;  CURRENT APPLICATION NUMBER: US/09/815,346
;  CURRENT FILING DATE:  2001-03-23
;  PRIOR APPLICATION NUMBER: 60/191,736
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 19
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-09-815-346-2

  Query Match             100.0%;  Score 41;  DB 9;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 5.9;
```

```
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          6 PDTRPAP 12


RESULT 15
US-09-815-346-6
; Sequence 6, Application US/09815346
; Publication No. US20020018806A1
; GENERAL INFORMATION:
;  APPLICANT: AGRAWAL, BABITA
;  APPLICANT:  LONGENECKER, MICHAEL B.
;  APPLICANT:  PARKER, JOANNE
;  TITLE OF INVENTION: LIPOPEPTIDE ADJUVANTS
;  FILE REFERENCE: 042881/0160
;  CURRENT APPLICATION NUMBER: US/09/815,346
;  CURRENT FILING DATE:  2001-03-23
;  PRIOR APPLICATION NUMBER: 60/191,736
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 19
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-09-815-346-6

  Query Match             100.0%;  Score 41;  DB 9;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 5.9;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          6 PDTRPAP 12


Search completed: December 29, 2004, 13:24:26
Job time : 422 secs
```

```
DR   WPI; 2002-083037/11.
XX
PT   New T cell binding ligand peptide for treating immunological disorders
PT   such as herpes simplex virus, tuberculosis, cancers, acquired
PT   immunodeficiency syndrome and allergies.
XX
PS   Disclosure; Page 26; 110pp; English.
XX
CC   The present invention relates to novel T-cell binding ligand (TCBL)
CC   peptides (e.g. peptide G' (modified human MHC class II beta chain peptide
CC   G, peptide J (human beta-2-microglobulin peptide) and HIV-1 peptides) and
CC   TCBL peptide constructs for treating immunological disorders. The peptide
CC   constructs are useful for eliciting a cellular immune response in a human
CC   patient. The method comprises administering the peptide construct to the
CC   patient preferably in combination with an immune response adjuvant. The
CC   peptide constructs in the form of conjugated peptides are useful for
CC   eliciting a cellular immune response in a patient exposed to or at risk
CC   for exposure to the human immunodeficiency virus (HIV). The TCBL peptides
CC   are useful for treating a patient suffering from an immunological
CC   disorder such as herpes simplex virus (HSV) infection, malaria,
CC   tuberculosis, cancers, acquired immunodeficiency syndrome (AIDS),
CC   allergies, autoimmune diseases (e.g. arthritis, Graves disease, multiple
CC   sclerosis (MS), autoimmune myocarditis, diabetes and lupus) by
CC   administering a peptide construct comprising a TCBL peptide bonded to an
CC   antigenic peptide associated with the disorder. Unlike prior art peptide
CC   conjugates, a modified version of peptide G has long range stabilisation
CC   and also enhances the immune response. AAU82019-AAU82114 represent T-cell
CC   specific binding ligand peptides, peptide constructs or peptides used in
CC   their construction
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 41;  DB 5;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          2 PDTRPAP 8


RESULT 5
AAE20933
ID   AAE20933 standard; peptide; 8 AA.
XX
AC   AAE20933;
XX
DT   01-JUL-2002  (first entry)
XX
DE   Human tumour-associated MUC1 antigen epitope #2.
XX
KW   Human; tumour-associated MUC1 antigen; epitope; cytostatic; vaccine;
KW   therapy; glycoprotein; immune system.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   3
FT                   /note= "Linked to single alpha-N-Acetyl-galactosamine
FT                   residue in O-glycosidic linkage"
XX
PN   EP1182210-A1.
XX
PD   27-FEB-2002.
XX
PF   17-AUG-2000; 2000EP-00117735.
XX
PR   17-AUG-2000; 2000EP-00117735.
XX
PA   (BAST/) BASTERT G.
XX
PI   Bastert G,  Kaul S;
XX
```

```
DR   WPI; 2002-282784/33.
XX
PT   Carbohydrate-containing epitopes of tumor-associated MUC1 antigens
PT   specifically bound by monoclonal antibody 7F11 and 1E4, useful as vaccine
PT   for treating humans e.g. tumor patient or animal.
XX
PS   Claim 3; Fig 4b; 38pp; English.
XX
CC   The invention relates to epitopes of tumour-associated mucin MUC1 antigen
CC   comprising an amino acid sequence with atleast one carbohydrate moiety at
CC   position 3. MUC1 is a transmembrane glycoprotein which is located on the
CC   luminal surface of ductal epithelial. A tumour vaccine and/or a
CC   pharmaceutical composition is useful for immunisation in humans (tumour
CC   patients) or animals and for the production of monoclonal antibodies. A
CC   diagnostic composition is useful in an diagnostic method useful for
CC   detection of tumour associated MUC1 antigen within body of a patient. A
CC   glycopeptide antigen is useful in a diagnostic method for the detection
CC   of antibodies and/or cells of the cellular or humoral immune system of an
CC   individual. The present sequence is human tumour-associated MUC1 antigen
CC   epitope
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 41;  DB 5;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          1 PDTRPAP 7


RESULT 6
ADL15165
ID   ADL15165 standard; peptide; 8 AA.
XX
AC   ADL15165;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Human tumour-associated antigen, MUC-1 peptide #2.
XX
KW   Therapeutic binding agent; tumour-associated antigen; MUC-1; cancer;
KW   tumour; breast carcinoma; colon carcinoma;
KW   oesophageal squamous cell carcinoma; pancreatic carcinoma;
KW   prostate carcinoma; multiple myeloma; cytostatic; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          3. .8
FT                   /note= "Specifically claimed in claim 1, and also given
FT                   as SEQ ID No:5"
XX
PN   US6716966-B1.
XX
PD   06-APR-2004.
XX
PF   18-AUG-2000; 2000US-00641833.
XX
PR   18-AUG-1999;   99US-0149492P.
PR   11-NOV-1999;   99US-0164714P.
XX
PA   (ALTA-) ALTAREX CORP.
XX
PI   Madiyalakan R;
XX
DR   WPI; 2004-303095/28.
XX
PT   New binding agent, Alt-1, that binds immunological determinants of MUC-1,
PT   useful for therapeutically treating a mammal bearing a tumor e.g. breast,
PT   colon, esophageal, prostate or pancreatic carcinoma, or multiple myeloma.
XX
```

```
PS   Claim 1; SEQ ID NO 2; 18pp; English.
XX
CC   The present invention relates to therapeutic binding agents that bind to
CC   tumour-associated antigen, MUC-1. The binding agent reduces, reverses, or
CC   prevents the effects of MUC-1 in cancer. The binding agent is useful for
CC   treating a mammal bearing a tumour e.g. breast carcinoma, colon
CC   carcinoma, oesophageal squamous cell carcinoma, pancreatic carcinoma,
CC   prostate carcinoma or multiple myeloma. The present sequence represents a
CC   human MUC-1 peptide.
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 41;  DB 8;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          2 PDTRPAP 8


RESULT 7
AAR68021
ID   AAR68021 standard; peptide; 9 AA.
XX
AC   AAR68021;
XX
DT   25-MAR-2003  (revised)
DT   05-SEP-1995  (first entry)
XX
DE   Mucin peptide p20-28.
XX
KW   Mucin; multiple tandem repeat; vaccine; pancreas cancer; breast cancer;
KW   ovary cancer; colon cancer; HIV.
XX
OS   Synthetic.
XX
PN   WO9503825-A1.
XX
PD   09-FEB-1995.
XX
PF   29-JUL-1994;   94WO-US008477.
XX
PR   30-JUL-1993;   93US-00099354.
XX
PA   (FINN/) FINN O J.
PA   (FONT/) FONTENOT J D.
PA   (MONT/) MONTELARO R C.
XX
PI   Finn OJ,  Fontenot JD,  Montelaro RC;
XX
DR   WPI; 1995-082033/11.
XX
PT   Synthetic multiple tandem repeat mucin-1 peptides and analogues - have
PT   native conformation in the absence of glycosylation and are linked to
PT   epitopes; for vaccines and tests of cancer, viruses and bacteria.
XX
PS   Disclosure; Page 54; 125pp; English.
XX
CC   A synthetic peptide such as p105 (AAR68022) includes 5 tandem repeats of
CC   the mucin peptide given in AAR68004. The DTR motif, located between the
CC   first 2 prolines of each repeat, is the target of an anti-mucin immune
CC   response, and can be substituted by a sequence from a virus, tumor
CC   antigen or autoantigen. Drugs based on 9-amino acid portions (AAR68007-
CC   21) of the mucin peptide were developed. (Updated on 25-MAR-2003 to
CC   correct PN field.)
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 41;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          2 PDTRPAP 8


RESULT 8
AAR68007
ID   AAR68007 standard; peptide; 9 AA.
XX
AC   AAR68007;
XX
DT   25-MAR-2003  (revised)
DT   05-SEP-1995  (first entry)
XX
DE   Mucin peptide p1-9.
XX
KW   Mucin; multiple tandem repeat; vaccine; pancreas cancer; breast cancer;
KW   ovary cancer; colon cancer; HIV.
XX
OS   Synthetic.
XX
PN   WO9503825-A1.
XX
PD   09-FEB-1995.
XX
PF   29-JUL-1994;   94WO-US008477.
XX
PR   30-JUL-1993;   93US-00099354.
XX
PA   (FINN/) FINN O J.
PA   (FONT/) FONTENOT J D.
PA   (MONT/) MONTELARO R C.
XX
PI   Finn OJ,  Fontenot JD,  Montelaro RC;
XX
DR   WPI; 1995-082033/11.
XX
PT   Synthetic multiple tandem repeat mucin-1 peptides and analogues - have
PT   native conformation in the absence of glycosylation and are linked to
PT   epitopes; for vaccines and tests of cancer, viruses and bacteria.
XX
PS   Disclosure; Page 54; 125pp; English.
XX
CC   A synthetic peptide such as p105 (AAR68022) includes 5 tandem repeats of
CC   the mucin peptide given in AAR68004. The DTR motif, located between the
CC   first 2 prolines of each repeat, is the target of an anti-mucin immune
CC   response, and can be substituted by a sequence from a virus, tumor
CC   antigen or autoantigen. Drugs based on 9-amino acid portions (AAR68007-
CC   21) of the mucin peptide were developed. (Updated on 25-MAR-2003 to
CC   correct PN field.)
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 41;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          1 PDTRPAP 7


RESULT 9
AAW72723
ID   AAW72723 standard; peptide; 9 AA.
XX
AC   AAW72723;
XX
DT   11-JAN-1999  (first entry)
XX
DE   Mucin peptide preparation p 20-28.
XX
KW   Mucin; muc-1; human; cancer; infectious disease; vaccine; diagnosis;
```

```
KW   multiple tandem repeat; pancreatic cancer; breast cancer; colon cancer.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   US5827666-A.
XX
PD   27-OCT-1998.
XX
PF   10-AUG-1994;   94US-00288059.
XX
PR   30-JUL-1993;   93US-00099354.
XX
PA   (UYPI-) UNIV PITTSBURGH.
XX
PI   Montelaro RC,  Fontenot JD,  Finn OJ;
XX
DR   WPI; 1998-593988/50.
XX
PT   Assay for cancer antibodies - using synthetic peptide comprising multiple
PT   tandem repeats of muc-1.
XX
PS   Disclosure; Col 25; 45pp; English.
XX
CC   An assay has been developed for antibodies to pancreatic, breast or colon
CC   cancer in a sample. The assay comprises contacting the sample with a
CC   synthetic muc-1 peptide that comprises at least two 20 amino acid tandem
CC   repeats of muc-1 and is capable of attaining native conformation in the
CC   absence of glycosylation, and detecting any peptide-antibody complex
CC   formation. The assay can be used in the diagnosis of e.g. pancreatic,
CC   breast or colon cancer. The present sequence represents a mucin peptide
CC   preparation from the present invention
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 41;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          2 PDTRPAP 8


RESULT 10
AAW72709
ID   AAW72709 standard; peptide; 9 AA.
XX
AC   AAW72709;
XX
DT   11-JAN-1999  (first entry)
XX
DE   Mucin peptide preparation p 1-9.
XX
KW   Mucin; muc-1; human; cancer; infectious disease; vaccine; diagnosis;
KW   multiple tandem repeat; pancreatic cancer; breast cancer; colon cancer.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   US5827666-A.
XX
PD   27-OCT-1998.
XX
PF   10-AUG-1994;   94US-00288059.
XX
PR   30-JUL-1993;   93US-00099354.
XX
PA   (UYPI-) UNIV PITTSBURGH.
XX
PI   Montelaro RC,  Fontenot JD,  Finn OJ;
XX
DR   WPI; 1998-593988/50.
XX
PT   Assay for cancer antibodies - using synthetic peptide comprising multiple
PT   tandem repeats of muc-1.
XX
PS   Disclosure; Col 25; 45pp; English.
XX
CC   An assay has been developed for antibodies to pancreatic, breast or colon
CC   cancer in a sample. The assay comprises contacting the sample with a
CC   synthetic muc-1 peptide that comprises at least two 20 amino acid tandem
CC   repeats of muc-1 and is capable of attaining native conformation in the
CC   absence of glycosylation, and detecting any peptide-antibody complex
CC   formation. The assay can be used in the diagnosis of e.g. pancreatic,
CC   breast or colon cancer. The present sequence represents a mucin peptide
CC   preparation from the present invention
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 41;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          1 PDTRPAP 7


RESULT 11
AAY05217
ID   AAY05217 standard; peptide; 9 AA.
XX
AC   AAY05217;
XX
DT   17-JUN-1999  (first entry)
XX
DE   MUC VNTR peptide SEQ ID NO 8.
XX
KW   VNTR peptide; mannose receptor; immunoregulatory composition; mucin;
KW   immune response; tumour; adenocarcinoma; breast cancer; therapy.
XX
OS   Homo sapiens.
XX
PN   WO9916455-A1.
XX
PD   08-APR-1999.
XX
PF   29-SEP-1998;   98WO-IB001718.
XX
PR   29-SEP-1997;   97US-0060594P.
XX
PA   (AUST-) AUSTIN RES INST.
XX
PI   Mckenzie IFC,  Apostolopoulos V,  Pietersz GA;
XX
DR   WPI; 1999-254927/21.
XX
PT   Immunoregulatory composition comprising mannose receptor-bearing cells,
PT   an antigen and mannose, useful for prevention/treatment of cancer.
XX
PS   Disclosure; Page 18; 84pp; English.
XX
CC   This sequence represents a VNTR from a form of human mucin. The invention
CC   relates to an immunoregulatory composition (I) comprising isolated
CC   mannose receptor-bearing cells and a conjugate comprising an antigen and
CC   a mannose selected from fully oxidised and partially reduced mannose
CC   having aldehydes. (I) is useful as a therapeutic agent in animals with
CC   natural antibodies against mucin, for inducing a cell mediated immune
CC   response specifically to mucin. It is useful for prevention/treatment of
CC   tumours, particularly adenocarcinoma, more particularly breast cancer.
CC   Prior art methods for removal of tumours uses disfiguring and costly
CC   surgery, and/or chemotherapeutic and radiation methods with severe side-
CC   effects. The new composition is non-toxic
XX
SQ   Sequence 9 AA;
```

```
  Query Match             100.0%;  Score 41;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          2 PDTRPAP 8


RESULT 12
AAW67597
ID   AAW67597 standard; peptide; 9 AA.
XX
AC   AAW67597;
XX
DT   02-MAR-1999  (first entry)
XX
DE   T-cell activation peptide #15.
XX
KW   Activated T helper cell; CD4+; cytotoxic T cell; CD8+; liposome; epitope;
KW   peripheral blood lymphocyte; antigen-presenting cell; APC; virus; tumour;
KW   bacterium; parasite; cytokine; vaccine; cancer; malaria; HIV; hepatitis;
KW   tuberculosis.
XX
OS   Synthetic.
XX
PN   WO9850527-A1.
XX
PD   12-NOV-1998.
XX
PF   07-MAY-1998;   98WO-US009288.
XX
PR   08-MAY-1997;   97US-0045949P.
XX
PA   (BIOM-) BIOMIRA INC.
XX
PI   Agrawal B,  Krantz MJ,  Reddish MA,  Longenecker BM;
XX
DR   WPI; 1999-034715/03.
XX
PT   Method of activation of T cells - by exposure to antigen-presenting cells
PT   loaded with antigen in liposome, used for, e.g. treating cancer and
PT   microbial infections.
XX
PS   Claim 14; Page 49; 75pp; English.
XX
CC   Peptides AAW67583-W67611 are used to produce activated T helper (CD4+)
CC   and cytotoxic (CD8+) T-cells. The activated T cells are produced by
CC   treating peripheral blood lymphocytes with liposome-encapsulated peptide
CC   antigen to generate Ag-loaded antigen-presenting cells (APC), contacting
CC   naive or anergic T-cells with these APC, and isolating the resulting
CC   activated T-cells. The cells are specific for a particular antigen,
CC   particularly one derived from a tumour, but also those from viruses,
CC   bacteria and other parasites. It can also be used to identify antigens
CC   and epitopes able to generate an Ag-specific T-cell response (by
CC   assessing proliferation and cytokine release). Also the Ag-loaded APC can
CC   be used as cellular vaccines for treating cancer (claimed) or other
CC   diseases (e.g. malaria, human immune deficiency virus infection,
CC   hepatitis, tuberculosis). The activated T-cells can be used to treat the
CC   same conditions by adoptive T-cell transfer therapy
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 41;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          1 PDTRPAP 7
```

```
RESULT 13
AAW67584
ID   AAW67584 standard; peptide; 9 AA.
XX
AC   AAW67584;
XX
DT   02-MAR-1999  (first entry)
XX
DE   T-cell activation peptide #2.
XX
KW   Activated T helper cell; CD4+; cytotoxic T cell; CD8+; liposome; epitope;
KW   peripheral blood lymphocyte; antigen-presenting cell; APC; virus; tumour;
KW   bacterium; parasite; cytokine; vaccine; cancer; malaria; HIV; hepatitis;
KW   tuberculosis.
XX
OS   Synthetic.
XX
PN   WO9850527-A1.
XX
PD   12-NOV-1998.
XX
PF   07-MAY-1998;   98WO-US009288.
XX
PR   08-MAY-1997;   97US-0045949P.
XX
PA   (BIOM-) BIOMIRA INC.
XX
PI   Agrawal B,  Krantz MJ,  Reddish MA,  Longenecker BM;
XX
DR   WPI; 1999-034715/03.
XX
PT   Method of activation of T cells - by exposure to antigen-presenting cells
PT   loaded with antigen in liposome, used for, e.g. treating cancer and
PT   microbial infections.
XX
PS   Claim 14; Page 49; 75pp; English.
XX
CC   Peptides AAW67583-W67611 are used to produce activated T helper (CD4+)
CC   and cytotoxic (CD8+) T-cells. The activated T cells are produced by
CC   treating peripheral blood lymphocytes with liposome-encapsulated peptide
CC   antigen to generate Ag-loaded antigen-presenting cells (APC), contacting
CC   naive or anergic T-cells with these APC, and isolating the resulting
CC   activated T-cells. The cells are specific for a particular antigen,
CC   particularly one derived from a tumour, but also those from viruses,
CC   bacteria and other parasites. It can also be used to identify antigens
CC   and epitopes able to generate an Ag-specific T-cell response (by
CC   assessing proliferation and cytokine release). Also the Ag-loaded APC can
CC   be used as cellular vaccines for treating cancer (claimed) or other
CC   diseases (e.g. malaria, human immune deficiency virus infection,
CC   hepatitis, tuberculosis). The activated T-cells can be used to treat the
CC   same conditions by adoptive T-cell transfer therapy
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 41;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          3 PDTRPAP 9


RESULT 14
AAW67596
ID   AAW67596 standard; peptide; 9 AA.
XX
AC   AAW67596;
XX
DT   02-MAR-1999  (first entry)
XX
DE   T-cell activation peptide #14.
XX
```

```
KW   Activated T helper cell; CD4+; cytotoxic T cell; CD8+; liposome; epitope;
KW   peripheral blood lymphocyte; antigen-presenting cell; APC; virus; tumour;
KW   bacterium; parasite; cytokine; vaccine; cancer; malaria; HIV; hepatitis;
KW   tuberculosis.
XX
OS   Synthetic.
XX
PN   WO9850527-A1.
XX
PD   12-NOV-1998.
XX
PF   07-MAY-1998;   98WO-US009288.
XX
PR   08-MAY-1997;   97US-0045949P.
XX
PA   (BIOM-) BIOMIRA INC.
XX
PI   Agrawal B,  Krantz MJ,  Reddish MA,  Longenecker BM;
XX
DR   WPI; 1999-034715/03.
XX
PT   Method of activation of T cells - by exposure to antigen-presenting cells
PT   loaded with antigen in liposome, used for, e.g. treating cancer and
PT   microbial infections.
XX
PS   Claim 14; Page 49; 75pp; English.
XX
CC   Peptides AAW67583-W67611 are used to produce activated T helper (CD4+)
CC   and cytotoxic (CD8+) T-cells. The activated T cells are produced by
CC   treating peripheral blood lymphocytes with liposome-encapsulated peptide
CC   antigen to generate Ag-loaded antigen-presenting cells (APC), contacting
CC   naive or anergic T-cells with these APC, and isolating the resulting
CC   activated T-cells. The cells are specific for a particular antigen,
CC   particularly one derived from a tumour, but also those from viruses,
CC   bacteria and other parasites. It can also be used to identify antigens
CC   and epitopes able to generate an Ag-specific T-cell response (by
CC   assessing proliferation and cytokine release). Also the Ag-loaded APC can
CC   be used as cellular vaccines for treating cancer (claimed) or other
CC   diseases (e.g. malaria, human immune deficiency virus infection,
CC   hepatitis, tuberculosis). The activated T-cells can be used to treat the
CC   same conditions by adoptive T-cell transfer therapy
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 41;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          2 PDTRPAP 8


RESULT 15
AAB72499
ID   AAB72499 standard; peptide; 9 AA.
XX
AC   AAB72499;
XX
DT   09-MAY-2001  (first entry)
XX
DE   Fusion protein peptide fragment #33.
XX
KW   Immunomodulatory; human milk fat globule specific antibody; HMFG; cancer;
KW   tumour detection.
XX
OS   Unidentified.
XX
PN   US6190885-B1.
XX
PD   20-FEB-2001.
XX
PF   08-OCT-1993;   93US-00134198.
XX
PR   02-FEB-1990;   90US-00473673.
PR   08-APR-1993;   93US-00046103.
XX
PA   (CANC-) CANCER RES FUND CONTRA COSTA.
XX
PI   Ceriani RL,  Peterson JA,  Larocca DJ;
XX
DR   WPI; 2001-217896/22.
XX
PT   Novel fusion protein comprising sequence of amino acids which binds
PT   antibodies specific to human milk fat globule differentiation antigens,
PT   useful as immunogen and for diagnosing breast cancers.
XX
PS   Disclosure; Col 19; 37pp; English.
XX
CC   The present invention relates to a fusion protein (see AAB72464) which
CC   substantially fails to bind to human milk fat globule (HMFG) specific
CC   antibody. The fusion protein can be used as an immunogen and for
CC   diagnostic purposes and as part of a kit for detecting the presence of
CC   neoplastic tissue from a solid tumour or metastasis. The present sequence
CC   is a peptide which was used in the present invention
XX
SQ   Sequence 9 AA;

  Query Match             100.0%;  Score 41;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PDTRPAP 7
              |||||||
Db          2 PDTRPAP 8


Search completed: December 29, 2004, 13:12:31
Job time : 78 secs
```